



Lehrstuhl für
Wirtschaftsinformatik
Information Systems
Management

No. 24

February 2007

Bayreuther Arbeitspapiere zur Wirtschaftsinformatik

Torsten Eymann (Hrsg.)

Tagungsband zum Doctoral Consortium der WI 2007

Bayreuth Reports on Information Systems Management



UNIVERSITÄT
BAYREUTH

ISSN 1864-9300

Die Arbeitspapiere des Lehrstuhls für Wirtschaftsinformatik dienen der Darstellung vorläufiger Ergebnisse, die i. d. R. noch für spätere Veröffentlichungen überarbeitet werden. Die Autoren sind deshalb für kritische Hinweise dankbar.

Alle Rechte vorbehalten. Insbesondere die der Übersetzung, des Nachdruckes, des Vortrags, der Entnahme von Abbildungen und Tabellen – auch bei nur auszugsweiser Verwertung.

Authors:

Torsten Eymann (University of Bayreuth)
et al.

The Bayreuth Reports on Information Systems Management comprise preliminary results which will usually be revised for subsequent publications. Critical comments would be appreciated by the authors.

All rights reserved. No part of this report may be reproduced by any means, or translated.

**Information Systems Management
Working Paper Series**

Edited by:

Prof. Dr. Torsten Eymann

Managing Assistant and Contact:

Raimund Matros
Universität Bayreuth
Lehrstuhl für Wirtschaftsinformatik (BWL VII)
Prof. Dr. Torsten Eymann
Universitätsstrasse 30
95447 Bayreuth
Germany

Email: raimund.matros@uni-bayreuth.de

ISSN 1864-9300

Inhaltsverzeichnis

Vorwort	5
Program committee	6
Programm	7
Entwicklung einer Methode zur Verbesserung der Entwurfsqualität von Serviceorientierten Architekturen auf Basis von Axiomatic Design, <i>René Fiege</i>	11
Modellierung und Messung des Einflusses von RFID-Technologie auf die Produktivität, <i>Daniel Gille</i>	29
Softwareagenten in der Krankenhauslogistik – Ein Ansatz zur effizienten Ressourcenallokation, <i>Christoph Niemann</i>	49
Kontextbasierte Adaption klinischer Pfade, <i>Claudia Reuter</i>	63
Simulation von Grid-Versicherungen - Eine Analyse der Risikoallokation in Grid-Netzwerken, <i>Werner Streitberger</i>	83
Incentive Engineering and Transaction Pricing in the Securities Trading Value Chain, <i>Matthias Burghardt</i>	99
Outsourcing Risk Mitigation - A Study of Business Process Outsourcing arrangements in the German banking industry, <i>Kim Wüllenweber</i>	115
Conception of XBRL Use in the Financial Reporting Chain, <i>Maciej Piechocki</i>	133
Potenzial experimenteller Forschung in der Wirtschaftsinformatik, <i>Thomas Wilde</i>	151
Optimale Preisgestaltung von internetbasierten Diensten, <i>Christian Schlereth</i>	169
Gestaltung kooperativer e-Financial Supply Chains, <i>Thomas Krabichler</i>	187
A Field Experiment on Monetary Incentives in Prediction Markets, <i>Stefan Luckner</i>	207
Zum Einsatz selbstorganisierender Koordinationsmechanismen unter Nutzung von UC-Technologien, <i>Titus Faupel</i>	219

Geschäftliche Verträge als Quelle für Services in unternehmensübergreifenden Informationssystemen, <i>Christian Brandt</i>	233
Indikatoren für Einsatz und Evaluation von Multiagentensystem-Technologien zur effizienten Steuerung von Organisationsabläufen in Krankenhäusern - am Beispiel eines Prozess in einem Krankenhaus in Deutschland und Indonesien, <i>Dwi Anoraganingrum</i>	247
Event analysis in Virtual Stock Markets for business research, <i>Christian Slamka</i>	263
Kosten-Nutzen-Quantifizierung von Investitionen in IT-Sicherheit als Grundlage eines ökonomisch fundierten Sicherheitsmanagements, <i>Thomas Nowey</i>	281
Analyse der Selbstorganisation in Wikis, <i>Claudia Müller</i>	299
Cannibalization Effects in Online & Offline Retailing, <i>Gottfried Gruber</i>	317
Ubiquitous Computing im Krankenhaus - eine fallstudienbasierte Untersuchung betriebswirtschaftlicher Potenziale und kritischer Erfolgsfaktoren, <i>Falk Zwicker</i>	335
Autoren	351

Vorwort

Die Wirtschaftsinformatik sieht sich im deutschsprachigen Raum einer zunehmenden Themen- und Methodenvielfalt sowie kürzeren Innovationszyklen ausgesetzt. Um diese Herausforderungen bewältigen zu können, kommt dem Kennenlernen und der Förderung des wissenschaftlichen Nachwuchses eine große Bedeutung zu. Daher wird bereits seit Anfang der 1990er Jahre jungen Wissenschaftlern im Vorfeld der Tagung "Wirtschaftsinformatik" ein Consortium als unterstützendes Forum angeboten. Damit werden folgende Ziele verfolgt:

- Vorstellung junger Wissenschaftler: Es stellen sich ambitionierte Nachwuchswissenschaftler vor, die mit einem Vortrag über ihr Dissertationsvorhaben ihre Arbeit präsentieren und ihre Fähigkeit zum wissenschaftlichen Diskurs einüben sollen.
- Unterstützung junger Wissenschaftler: Der wissenschaftliche Nachwuchs soll eine inhaltliche und methodische Unterstützung bei der Fortführung seiner Arbeiten durch erfahrene Wissenschaftler erhalten.
- Community Building: Die Veranstaltung dient dem Kennenlernen zwischen Professoren und Doktoranden. Dies soll den Diskurs und den Zusammenhalt in der Disziplin erhöhen und jungen Wissenschaftlern den Zugang zur "Scientific Community" erleichtern.

Insofern richtet sich das Doctoral Consortium an alle Nachwuchswissenschaftler der Wirtschaftsinformatik, die eine fundierte Dissertation anstreben und den wissenschaftlichen Austausch intensivieren möchten.

Das Doctoral Consortium 2007 fand an zwei Tagen vor der eigentlichen WI-Konferenz, von Montag, 26.2. bis Dienstag, 27.2. in Karlsruhe statt. Aufgabe der Mentoren war es, dem wissenschaftlichen Nachwuchs eine inhaltliche und methodische Unterstützung bei der Fortführung seiner Arbeiten durch erfahrene Wissenschaftler zu geben. Dazu stellten sich ambitionierte Nachwuchswissenschaftler vor, die mit einem Vortrag über ihr Dissertationsvorhaben ihre Arbeit präsentieren und ihre Fähigkeit zum wissenschaftlichen Diskurs einüben sollen. In kleinen Arbeitsgruppen (ca. 2 Professoren und 4-5 Doktoranden) wurden die Vorhaben in den zwei Tagen vertieft diskutiert. Eingeladene Vorträge, eine von Sponsoren unterstützte Abendveranstaltung sowie weitere Elemente rundeten das Programm ab.

Als Mentoren bin ich folgenden Kolleginnen und Kollegen zu Dank verpflichtet:

Abraham Bernstein, Universität Zürich
Michael Breitner, Leibniz Universität Hannover
Hans-Ulrich Buhl, Universität Augsburg
Peter Buxmann, TU Darmstadt
Thomas Hess, Ludwig-Maximilian-Universität München
Helmut Krcmar, Technische Universität München
Günter Müller, Albert-Ludwigs-Universität Freiburg
Mareike Schoop, Universität Hohenheim
Myra Spiliopoulou, Otto-von-Guericke-Universität Magdeburg

Das Doctoral Consortium wurde von der Unternehmensberatung A.T. Kearney finanziell unterstützt. Hier danke ich in erster Linie Herrn Holger Röder.

Bayreuth, im Februar 2007

Torsten Eymann

Nachwuchsobmann der Wissenschaftlichen Kommission Wirtschaftsinformatik
im Verband der Hochschullehrer für Betriebswirtschaft e.V.

Program committee

Torsten Eymann, Universitaet Bayreuth, Deutschland	Armin Heinzl, Universität Mannheim
Abraham Bernstein, Universität Zürich	Thomas Hess, LMU München
Myra Spiliopoulou, Universität Magdeburg	Michael Amberg, Universität Erlangen-Nürnberg
Helmut Krcmar, TU München	Jörg Müller, TU Clausthal
Günter Müller, Universität Freiburg	Hans-Ulrich Buhl, Universität Augsburg
Michael Breitner, Universität Hannover	Peter Buxmann, TU Darmstadt
Mareike Schoop, Universität Hohenheim	

Programm

26. Februar 2007

14:00 : *Eröffnungsplenum*

14:15 : *Arbeitsgruppen*

14:30 : *Vortrag 1*

15:45 : *Vortrag 2*

17:15 : *Plenumsvortrag*

Next Generation Outsourcing - more value for the business?,
Röder, H.; Dr. Eul, M.; Kannegießer, M., Klapdor, S.

18:30 : *Abendessen auf Einladung von A.T. Kearney*

27. Februar 2007

9:00 : *Vortrag 3*

10:15 : *Vortrag 4*

14:00 : *Eingeladener Plenumsvortrag*

Ph. D. Studies and Academic Career in US, D, CH and A,
Prof. Dr. Avi Bernstein, Universität Zürich

15:00 : *Plenumsdiskussion*

Doktorandenausbildung und zukünftige Doktorandenprogramme - Meinungen und Sicht der Teilnehmer des DC

Arbeitsgruppe Prof. Dr. Breitner, Universität Hannover Prof. Dr. Krcmar, TU München :

Entwicklung einer Methode zur Verbesserung der Entwurfsqualität von Serviceorientierten Architekturen auf Basis von Axiomatic Design, <i>Rene Fiege</i>	11
Modellierung und Messung des Einflusses von RFID-Technologie auf die Produktivität, <i>Daniel Gille</i>	29
Softwareagenten in der Krankenhauslogistik – Ein Ansatz zur effizienten Ressourcenallokation, <i>Christoph Niemann</i>	49
Kontextbasierte Adaption klinischer Pfade, <i>Claudia Reuter</i>	63

Arbeitsgruppe Prof. Dr. Buhl, Universität Augsburg Prof. Dr. Eymann, Universität Bayreuth :

Simulation von Grid-Versicherungen - Eine Analyse der Risikoallokation in Grid-Netzwerken, <i>Werner Streitberger</i>	83
Incentive Engineering and Transaction Pricing in the Securities Trading Value Chain, <i>Matthias Burghardt</i>	99
Outsourcing Risk Mitigation - A Study of Business Process Outsourcing arrangements in the German banking industry, <i>Kim Wüllenweber</i>	115
Conception of XBRL Use in the Financial Reporting Chain, <i>Maciej Piechocki</i>	133

Arbeitsgruppe Prof. Dr. Bernstein, Universität Zürich Prof. Dr. Buxmann, TU Darmstadt :

Potenzial experimenteller Forschung in der Wirtschaftsinformatik, <i>Thomas Wilde</i>	151
Optimale Preisgestaltung von internetbasierten Diensten, <i>Christian Schlereth</i>	169
Gestaltung kooperativer e-Financial Supply Chains, <i>Thomas Krabichler</i>	187
A Field Experiment on Monetary Incentives in Prediction Markets, <i>Stefan Luckner</i>	207

Arbeitsgruppe Prof. Dr. Hess, LMU München Prof. Dr. Schoop, Universität Hohenheim :

Zum Einsatz selbstorganisierender Koordinationsmechanismen unter Nutzung von UC-Technologien, <i>Titus Faupel</i>	219
Geschäftliche Verträge als Quelle für Services in unternehmensübergreifenden Informationssystemen, <i>Christian Brandt</i>	233
Indikatoren für Einsatz und Evaluation von Multiagentensystem-Technologien zur effizienten Steuerung von Organisationsabläufen in Krankenhäusern - am Beispiel eines Prozess in einem Krankenhaus in Deutschland und Indonesien, <i>Dwi Anoraganingrum</i>	247
Event analysis in Virtual Stock Markets for business research, <i>Christian Slamka</i>	263

Arbeitsgruppe Prof. Dr. Müller, Universität Freiburg Prof. Dr. Spiliopolou, Universität Magdeburg :	
Kosten-Nutzen-Quantifizierung von Investitionen in IT-Sicherheit als Grundlage eines Ökonomisch fundierten Sicherheitsmanagements,	
<i>Thomas Nowey</i>	281
Analyse der Selbstorganisation in Wikis,	
<i>Claudia Müller</i>	299
Cannibalization Effects in Online & Offline Retailing,	
<i>Gottfried Gruber</i>	317
Ubiquitous Computing im Krankenhaus - eine fallstudienbasierte Untersuchung betriebswirtschaftlicher Potenziale und kritischer Erfolgsfaktoren,	
<i>Falk Zwicker</i>	335

Proposal für das Dissertationsthema

**Entwicklung einer Methode zur Verbesserung der
Entwurfsqualität von Serviceorientierten Architekturen auf Basis
von Axiomatic Design**

Betreuender Hochschullehrer

Univ.-Prof. Dr. rer. pol. habil. Dirk Stelzer

Fachgebiet Informations- und Wissensmanagement

Technische Universität Ilmenau

98693 Ilmenau

dirk.stelzer@tu-ilmenau.de

Bearbeiter des Themas

Dipl.-Wirtsch.-Inf. René Fiege

Fachgebiet Informations- und Wissensmanagement

Technische Universität Ilmenau

98693 Ilmenau

rene.fiege@tu-ilmenau.de

Abstract

Axiomatic Design (AD) ist eine Methode, die den Entwurf beliebiger Systeme unterstützen kann. AD hilft, Anforderungen klar voneinander abzugrenzen und unterstützt die Entwicklung von Systemen, deren Komponenten eine überschaubare Komplexität aufweisen und weitgehend unabhängig voneinander sind. Diese Ziele des AD korrespondieren mit wesentlichen Architekturzielen für Serviceorientierte Architekturen (SOA), insbesondere „ausgewogene Granularität“, „lose Kopplung“ und „hohe Autonomie“ von Services. Diese Ziele sind wichtig, um die Wiederverwendbarkeit und Komponierbarkeit der Services der SOA zu erhöhen. In dieser Dissertation wird untersucht, welchen Beitrag AD zum Entwurf von SOA leisten kann. Das Ziel dieser Arbeit besteht in der Entwicklung einer Methode, in deren Rahmen AD für den Entwurf von SOA nutzbar gemacht wird, um die Erreichung der Architekturziele von SOA zu verbessern. Darüber hinaus wird überprüft, inwiefern diese Methode im Entwurf von SOA dazu beiträgt, die Architekturziele von SOA – insbesondere lose Kopplung, hohe Autonomie, ausgewogene Granularität und Wiederverwendbarkeit – zu erreichen.

1 Einführung

1.1 Problemstellung

Serviceorientierte Architekturen (SOA) sollen es ermöglichen, wandlungsfähige bzw. „agile“ [HaSc2006, 276] Architekturen für Informationssysteme (IS) zu realisieren, so dass diese leicht an neue Anforderungen angepasst werden können [SiHu2005, 71 ff.]. In SOA werden so genannte Services einer Vielzahl von Teilnehmern zur Nutzung bereitgestellt. Services kapseln wiederverwendbare Funktionen. Sie sollen lose gekoppelt sein und je nach Bedarf zu beliebigen Anwendungen zusammengestellt werden können. Das SOA-Konzept ist relativ jung und befindet sich immer noch in einer Phase der Erprobung und Weiterentwicklung [Erl2005, 72]. Obwohl es bereits einige viel versprechende Modelle zur Unterstützung von Entwurf, Implementierung, Betrieb und Wartung serviceorientierter Systeme gibt [z. B. BuGa2005; EAAC2004, 83 ff.; Erl2005, 359 ff.; KoHB2005; LaMB2005; MaBe2006, 99-149], sind verschiedene Herausforderungen bisher nicht zufrieden stellend gelöst worden. Hierzu gehört unter anderem, wie Services mit einer ausgewogenen Granularität entworfen werden können und wie bereits im Entwurf darauf hingewirkt werden kann, dass Services entstehen, welche in sich möglichst autonom und untereinander lose gekoppelt sind. Die Architekturziele einer SOA, z. B. ausgewogene Granularität, lose Kopplung und hohe Autonomie der Services, sind wichtig, um die Wiederverwendbarkeit der Services zu erhöhen [Erl2005, 290 ff.]. Axiomatic Design (AD) ist eine Methode zur strukturierten Gestaltung von Objekten (z. B. Materialien, beliebige Systeme, etc.) [Suh2001]. Urheber und Anwender von AD behaupten [DoPa2001, 328; DoSu1999, 121; Jams2004, 1; SuDo2000, 95-96; Suh2001, 239 ff.; YiPa2004, 1, 6], dass diese Methode geeignet ist, Systeme zu entwerfen, deren Komponenten eine überschaubare Komplexität aufweisen sowie weitgehend unabhängig voneinander sind und dass Anforderungen an das zu entwerfende System klar voneinander abgegrenzt werden können [Suh2001, 29 ff.]. Diese Ziele des AD korrespondieren mit den Architekturzielen für SOA.

1.2 Zielsetzung

Das Ziel dieser Arbeit besteht in der Entwicklung einer Methode, in deren Rahmen AD für den Entwurf von SOA nutzbar gemacht wird, um die Erreichung der Architekturziele von SOA zu verbessern. Darüber hinaus wird überprüft, inwiefern diese Methode beim Entwurf von SOA dazu beiträgt, die Architekturziele von SOA – insbesondere lose Kopplung, hohe Autonomie, ausgewogene Granularität und Wiederverwendbarkeit – zu erreichen.

2 Grundlagen des Axiomatic Design

AD wurde Ende der 70er Jahre von Nam Pyo Suh am Massachusetts Institute of Technology (MIT) entwickelt [Suh1990, 18]. Es handelt sich dabei um eine Methode, mit der man strukturiert beliebige Objekte (z. B. Materialien, beliebige Systeme, etc.) entwerfen kann. Das Grundprinzip von AD umfasst die strukturierte Suche und Zuordnung geeigneter Lösungen für zuvor festgelegte Anforderungen. Ein Entwurf ist definiert als das Ergebnis dieses Zuordnungsprozesses [Suh2001, 2 ff.]. Er beschreibt, welche Anforderungen durch welche Lösung erfüllt werden können. AD basiert auf dem Konzept der Domänen, dem Dekompositionsprozess sowie auf dem so genannten Unabhängigkeits- und dem Informationsaxiom. Beide Axiome formulieren Richtlinien für den Entwurfsprozess.

2.1 Bestandteile von Axiomatic Design

Das Konzept der Domänen umfasst die Kundendomäne, die funktionale und die physische Domäne sowie die Prozessdomäne. Der Entwurfsprozess erstreckt sich über alle Domänen. Er beginnt in der Kundendomäne und endet in der Prozessdomäne. Jede Vorgängerdomäne beschreibt Anforderungen, jede Folgedomäne die korrespondierenden Lösungen. Zwischen allen Domänen erfolgt eine Zuordnung von Anforderungen zu korrespondierenden Lösungen [Suh2001, 10-14].

Während der Zuordnung zwischen den Domänen wird das Unabhängigkeitsaxiom angewendet. Im Folgenden wird dies exemplarisch anhand der Zuordnung zwischen der funktionalen und der physischen Domäne erläutert. Das Unabhängigkeitsaxiom verlangt, dass die Unabhängigkeit der so genannten funktionalen Anforderungen in der funktionalen Domäne nach Zuordnung so genannter Designparameter in der physischen Domäne gewahrt bleibt. Vollständige Unabhängigkeit liegt dann vor, wenn jede funktionale Anforderung durch genau einen Designparameter erfüllt wird. Der Zuordnungsprozess zwischen den Domänen wird hierarchisch im Top-down-Verfahren durchgeführt, um den Entwurf weiter zu verfeinern. Man spricht in diesem Zusammenhang vom Dekompositionsprozess (Abb. 1). Der Dekompositionsprozess verlangt, dass zwischen den Domänen hin und her gesprungen wird. Wie in Abb. 1 dargestellt, springt man ausgehend von einer funktionalen Anforderung in die physische Domäne, um einen geeigneten Designparameter zuzuordnen. Anschließend erfolgt der Rücksprung in die funktionale Domäne.

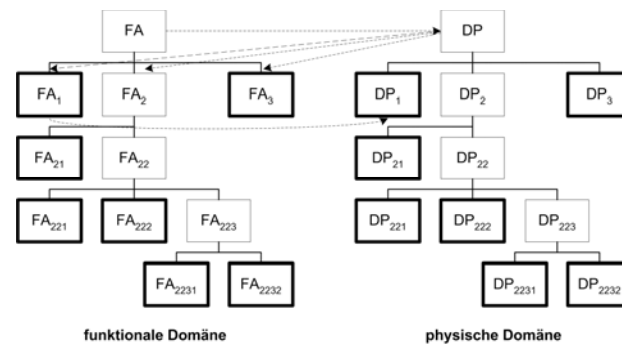


Abb. 1: Dekompositionsprozess

Dieser Prozess wird solange wiederholt, bis so genannte „elementare FA-DP-Kombinationen“ gefunden wurden. Eine Elementarkombination (diese sind in Abb. 1 fett hervorgehoben) liegt vor, wenn für eine funktionale Anforderung ein Designparameter gefunden wird, der unmittelbar, d. h. ohne weitere Dekomposition, implementierbar ist. Sobald eine Hierarchieebene im Dekompositionsprozess fertig gestellt wurde, wird das Unabhängigkeitsaxiom zur Prüfung der Unabhängigkeit der funktionalen Anforderungen herangezogen. Erst danach wird der Dekompositionsprozess auf tiefer liegenden Ebenen fortgesetzt.

Liegen alternative FA-DP-Kombinationen vor, die das Unabhängigkeitsaxiom erfüllen, wird das Informationsaxiom angewendet, um den besten Entwurf zu ermitteln. Das Informationsaxiom ermöglicht eine quantitative Bewertung gegebener Entwürfe und ermöglicht eine Reduktion der Komplexität sowie eine Erhöhung der Zuverlässigkeit [CIHi2000, 274; Suh2001, 39 ff.]. Das Informationsaxiom verlangt, dass der so genannte Informationsgehalt reduziert wird. Von den Entwürfen, die das Unabhängigkeitsaxiom erfüllen, wird derjenige ausgewählt, bei welchem der geringste Informationsgehalt ermittelt wird. Man spricht in diesem Zusammenhang vom Selektionsprozess.

2.2 Vorteile von Axiomatic Design

Zu den wesentlichen Merkmalen des AD gehören die Strukturierung von Entwurfsprozessen und die Reduzierung der Komplexität von Entwurfsergebnissen [Suh2001, 5 ff.]. AD bewirkt, dass während des Entwurfsprozesses nur für relevante Anforderungen Lösungen erarbeitet werden. Dadurch reduziert sich die Anzahl der Bestandteile eines Entwurfsobjektes [Suh2001, 29 ff.]. Entwurfsobjekte werden in hierarchisch strukturierte Bestandteile und Subbestandteile mit (a) überschaubarer Komplexität entworfen [Suh2001, 30]. Das Unabhängigkeitsaxiom sorgt

dafür, dass (b) die Unabhängigkeit der Objektbestandteile gewahrt bleibt und somit möglichst wenige Beziehungen zwischen den Bestandteilen entstehen. Das hierarchisch strukturierte Top-Down-Vorgehen sorgt darüber hinaus für (c) eine klare Abgrenzung der Anforderungen an die Entwurfsobjekte.

3 Architekturziele von Serviceorientierten Architekturen

Architekturziele repräsentieren Prinzipien, die eine SOA charakterisieren [Erl2005, 290]. Diese Ziele müssen bereits im Entwurf berücksichtigt werden, damit sie sich in der resultierenden Architektur widerspiegeln. Typische Architekturziele von SOA sind: Agilität, Wiederverwendbarkeit und Komponierbarkeit sowie angemessene Granularität, lose Kopplung, hohe Autonomie, Zustandlosigkeit, Auffindbarkeit, Abstraktheit, Interoperabilität, Geschäftsorientiertheit, Nachhaltigkeit, Neutralität und wohldefinierter Servicekontrakt [BuGa2005, 602; DJMZ2005, 9; EKAP2005, 28; Erl2005, 290; HaSc2006, 277; MaBe2006, 39 ff.; SiHu2005, 76-77]. Sowohl aus Sicht der Forschung als auch aus Sicht der Anwendung gehört zu den wichtigsten dieser Ziele die Erreichung einer hohen Wiederverwendbarkeit der Services [Berl2006, 4 ff.; EKAP2005; Erl2005, 290; Gall2004, 256]. Aus Sicht der Anwendung sollte sie möglichst hoch sein, damit die SOA an veränderte Anforderungen ohne unangemessen hohen Entwicklungsaufwand angepasst werden kann [Erl2005, 292]. Zeit- und Kosteneinsparungen lassen sich realisieren, weil derselbe Programmcode, der hinter einem Service gekapselt wurde, in verschiedenen Anwendungskontexten nicht erneut entwickelt werden muss, sondern einfach wieder verwendet werden kann [Berl2006, 4]. Aus Sicht der Forschung basiert der Grundgedanke der Serviceorientierung (wie in der Komponenten- oder Objektorientierung) auf dem im Software Engineering bekannten Prinzip der „separation of concerns“ [Balz1998; Erl2005, 290]. Dieses Grundprinzip geht auf Dijkstra zurück [Dijk1982]. Es beschreibt die Art und Weise der Dekomposition eines Problems in viele individuell voneinander identifizierbare Subprobleme. Ziel ist es, durch einen modularen Aufbau die Wiederverwendbarkeit einzelner Teile eines Softwaresystems zu gewährleisten. Ausgehend vom Grundprinzip der „separation of concerns“ hängt die Wiederverwendbarkeit der Services direkt von der Erfüllung der Ziele ausgewogene Granularität, lose Kopplung und hohe Autonomie ab [Balz1998; EKAP2005, 28; Erl2005, 290 ff.; HaSc2006, 281; MaBe2006, 39 ff.; PaYa2002, 58-60].

4 Stand der Forschung

In diesem Kapitel wird ein Überblick über den Stand der Forschung gegeben. Dies betrifft Vorgehensmodelle und Hilfsmittel, die den Entwurf von SOA unterstützen sollen. Außerdem werden die Anwendungsgebiete aufgezeigt, in welchen AD bisher erfolgreich eingesetzt wurde.

4.1 Vorgehensmodelle, Hilfsmittel im Entwurf von Serviceorientierten Architekturen

Die meisten Vorgehensmodelle zur Entwicklung von SOA unterscheiden die Phasen Entwurf, Implementierung, Test und Inbetriebnahme [z. B. BuGa2005; EAAC2004, 83 ff.; Erl2005, 359 ff.; KoHB2005; LaMB2005; MaBe2006, 99-149]. Tab. 1 zeigt Einzelheiten der Entwurfsphase einiger Vorgehensmodelle im Überblick.

Entwurf nach [EAAC2004, 83 ff.]	Entwurf nach [Erl2005, 358 ff.]	Entwurf nach [KoHB2005, 158 ff.]	Entwurf nach [MaBe2006, 99 ff.]
Identification <ul style="list-style-type: none"> - Domain decomposition/ Existing system analysis - Goal-service modeling Specification <ul style="list-style-type: none"> - Subsystem analysis - Component specification - Service allocation 	Service-oriented analysis <ul style="list-style-type: none"> - Define business automation requirements - Identify existing automation systems - Model candidate services Service-oriented Design <ul style="list-style-type: none"> - Compose SOA (layers, standards, extensions) - Design services - Design service-oriented business process 	Define system requirements <ul style="list-style-type: none"> - Elicit requirements - Rank requirements - Model requirements Design system architecture <ul style="list-style-type: none"> - Partition Services into abstract sub-systems - Establish sub-system interfaces 	Service Analysis and Identification Process <ul style="list-style-type: none"> - Discover Conceptual Business Services - Derive Candidate Business Services - Build Granularity Map - Apply Logical Operations on Candidate Business Services - Derive Actual Business Services Service Design Process <ul style="list-style-type: none"> - Examine Business Services State - Build Business Services Granularity Maps - Build Demarcation Maps - Apply Design Operations on Business Services - Realize Solution Services

Tab. 1: Entwurf Serviceorientierter Architekturen

Aus Tab. 1 geht hervor, dass die meisten Vorgehensmodelle den Entwurf in die Phasen Serviceorientierte Analyse und Design unterteilen. Den Modellen ist gemein, dass im Entwurf ausgehend von den Anforderungen an eine zu entwickelnde SOA das Gesamtsystem rekursiv in immer feiner granulierte Subsysteme dekomponiert wird (Prinzip der „separation of concerns“). Als Ergebnis dieses Dekompositionsprozesses entsteht ein Entwurfskonzept für eine SOA, das einzelne Servicekandidaten und deren Komposition zu höherwertigen Services beschreibt. Bisher gibt es nur wenige Hilfsmittel, um diesen Dekompositionsprozess strukturiert zu unterstützen. Zum Beispiel helfen die „Granularity Matrix“ und die „Granularity Map“, einzelne Services in Abhängigkeit von ihrem Abstraktionsniveau auf die richtige Hierarchieebene einer Dekompositionsstruktur einzuordnen [MaBe2006, 111, 124]. Die so genannte „Demarcation Map“ hilft ebenfalls bei dieser Einordnung. Auf ihrer Basis können außerdem Entscheidungen zur Aufspaltung (Dekomposition) oder Konsolidierung einzelner Services getroffen werden [MaBe2006, 126]. Kompositionsdiagramme helfen, einzelne

Services und deren Kompositionsbeziehungen übersichtlich in einer Baumstruktur darzustellen [Erl2005, 443]. Im Rahmen der so genannten Domänendekomposition wird ein Vorschlag für einzelne Services und deren Anordnung in einer Dekompositionshierarchie entwickelt [EAAC2004, 85]. Dabei geht man von der Zerlegung der Geschäftsprozesse aus, die durch die SOA unterstützt werden sollen. Die aufgeführten Vorgehensmodelle und Hilfsmittel bieten zwar viel versprechende Lösungsansätze für den Entwurf von SOA. Allerdings mangelt es ihnen an einer hinreichend feinen Struktur, die einem Entwickler detaillierte Schritte oder Regeln vorgibt, um den Dekompositionsprozess im Entwurf von SOA durchzuführen [KoHB2005, 157]. Gerade dieser Dekompositionsprozess ist aber wichtig, um die Architekturziele von SOA (vgl. Abschnitt 3) zu erreichen [KoHB2005, 157; ZSWP2005, 607]. Auf Grund der mangelnden Struktur können Entwurfsentscheidungen nur subjektiv getroffen werden, ohne explizit und bewusst auf bestimmte Architekturziele von SOA hinzuwirken [BuGa2005, 602; EKAP2005; Gall2004, 236 ff.].

4.2 Anwendungsgebiete des Axiomatic Design

AD wurde für den Maschinenbau entwickelt und zunächst zum Entwurf von Produkten angewendet [Suh2001, 11]. Mittlerweile wurde diese Methode auch in vielen anderen Gebieten erfolgreich eingesetzt. Tab. 2 gibt einen Überblick über diese Anwendungsgebiete und zeigt exemplarisch jeweils eine Quelle, die den Einsatz von Axiomatic Design beschreibt.

Gebiet	Entwurfsgegenstand	Quelle
Maschinenbau	Entwurf eines Produktionssystems	[CoRe1996]
Materialentwicklung	Entwurf neuartiger Materialien bestehend aus mikrozellularem Kunststoff	[Suh2001, 352 ff.]
Organisationsentwurf	Entwurf einer Organisationsstruktur	[WKAd2000]
Produktentwicklung	Entwurf einer Fahrzeugaufhängung	[DeSu2004]
Projektmanagement	Entwurf eines Projektplanes	[StDu2000]
Softwareentwicklung	Entwurf objektorientierter Software	[SuDo2000]
Strategische Planung	Entwurf eines Geschäftsplanes	[EnNo2000]
Systemtechnik	Entwurf beliebiger Systeme bestehend aus Hard- und Softwarekomponenten	[Suh1998]

Tab. 2: Anwendungsgebiete des Axiomatic Design

Aus dieser Übersicht wird deutlich, dass AD bisher noch nicht im Entwurf von SOA angewendet wurde. Die Beispiele in Tab. 2 belegen, dass die Vorteile des AD (vgl. Abschnitt 2.2) in den Entwurfsprojekten tatsächlich erzielt werden konnten.

5 Begründung der Forschung

Serviceorientierung und Architekturforschung sind aktuelle und relevante Forschungsgebiete der Wirtschaftsinformatik [BHMS2005; DJMZ2005; Mert2006; SSLD2005]. Aus Sicht der Forschung repräsentiert die Serviceorientierung neben der Komponenten- oder Objektorientierung ein neues Paradigma, um wiederverwendbare IS Architekturen zu schaffen [KoHB2005]. Die Fülle¹ wissenschaftlicher Abhandlungen und die große Anzahl von Konferenzen zu diesem Thema untermauern die Bedeutung aus Sicht der Forschung. SOA haben ein hohes Anwendungspotential in der Praxis. Hierfür spricht, dass viele große Softwareunternehmen, z. B. IBM, SAP, Software AG, Bea, Sun Microsystems, Oracle oder Microsoft, Softwarelösungen für oder auf Basis von SOA anbieten [BiLi2006, 99]. Viele Unternehmen beginnen zurzeit damit, ihre bestehenden IS Architekturen auf SOA umzustellen [Beri2006]. Ursache hierfür sind die Vorteile, die sich in den Architekturzielen von SOA widerspiegeln [Erl2005, 290 ff.]. Die Serviceorientierung ist zwar mit der Komponenten- oder Objektorientierung verwandt, dennoch eignen sich die vorhandenen Vorgehensmodelle und Hilfsmittel nur bedingt für den Entwurf von SOA [Gall2004, 236; KoHB2005; ZSWP2005, 607]. Die Spezifika der Serviceorientierung (z. B. Erreichung einer ausgewogenen Granularität) erfordern die Entwicklung neuer, strukturierte Vorgehensmodelle und Hilfsmittel. Diese Aufgabe ist ein wichtiger Teilbereich der Forschung zur Serviceorientierung an dem derzeit intensiv gearbeitet wird [BuGa2005; EKAP2005; Gall2004; KoHB2005; ZSWP2005].

5.1 Herausforderungen und Defizite im Entwurf von Serviceorientierten Architekturen

Die Ausführungen des Abschnittes 3 machen deutlich, dass bereits die Entwurfsphase eines Entwicklungsprojektes von SOA einen entscheidenden Einfluss auf die spätere Wiederverwendbarkeit einzelner Services hat. Herausforderungen bestehen insbesondere in der Erreichung der Architekturziele angemessene Granularität, lose Kopplung und hohe Autonomie [Balz1998; EKAP2005, 28; Erl2005, 290 ff.; HaSc2006, 281; MaBe2006, 39 ff.; PaYa2002, 58-60]. Abschnitt 4.1 zeigt, dass es bereits einige viel versprechende Vorgehensmodelle und Hilfsmittel zur Unterstützung des Entwurfs serviceorientierter Systeme gibt. Allerdings werden durch sie die Erreichung der Architekturziele bisher nicht zufrieden stellend gelöst. Die meisten Vorgehensmodelle und Hilfsmittel sind nicht detailliert genug, um ausgewählte Architekturziele

¹ Die Suchanfrage „SOA“ im Verbundkatalog GVK-PLUS (<http://gso.gbv.de/>) ergab 2.980 Treffer (Abruf am 2006-11-30).

bewusst und strukturiert anzustreben [BuGa2005, 602; EKAP2005; Gall2004, 236 ff.]. Im Vorfeld dieser Dissertation wurden Gespräche mit Experten bei Deutsche Post AG, Bea Systems GmbH, IBM, Quimonda AG, Sun Microsystems und Software AG geführt.² Deren Ergebnis bestätigt die identifizierten Herausforderungen und Defizite in der täglichen Praxis. Diese Ausführungen untermauern den Bedarf zur Entwicklung einer Methode, die dabei hilft, den Entwurf von SOA unter Berücksichtigung der Architekturziele von SOA besser zu strukturieren.

5.2 Axiomatic Design im Entwurf von Serviceorientierten Architekturen

AD wurde bisher noch nicht im Entwurf von SOA eingesetzt (vgl. Abschnitt 4.2). Es gibt zwei grundlegende Gemeinsamkeiten zwischen AD und dem Entwurf von SOA: 1. Sowohl im AD als auch im Entwurf von SOA findet eine hierarchische Zerlegung statt. In der Entwurfsphase der Vorgehensmodelle im Abschnitt 4.1 wird ausgehend von den Anforderungen an die zu entwickelnde SOA das Gesamtsystem rekursiv in immer feiner granulierte Subsysteme dekomponiert (Prinzip der „separation of concerns“). Die hierarchische Zerlegung ist eine wesentliche Grundlage, auf der AD beruht. Sie findet während des Dekompositionsprozesses statt (vgl. Abschnitt 2.1). 2. Wie die Untersuchungen des Forschungsprojektes „AD-SOA“ am Fachgebiet Informations- und Wissensmanagement der Technischen Universität Ilmenau zeigen, korrespondieren die Vorteile des AD mit den Architekturzielen von SOA [FiSt2007]. Beispielsweise führt die Strukturierung von Entwurfsobjekten in Bestandteile überschaubarer Komplexität zur Kreierung von Services in ausgewogener Granularität. Da AD die Unabhängigkeit der Objektbestandteile anstrebt, werden lose gekoppelte Services entworfen. Das hierarchische Top-Down-Vorgehen im AD bewirkt, dass Anforderungen an die SOA klar voneinander abgegrenzt werden. Dies führt zum Entwurf von autonomen Services, die klar nach den Aufgaben, die sie erfüllen, differenziert werden können.

Diese Gemeinsamkeiten zeigen, dass AD prinzipiell für eine Anwendung im Entwurf von SOA geeignet ist. Die Eigenschaften von AD adressieren die Herausforderungen und Defizite im Entwurf von SOA. AD könnte daher zu einer Verbesserung des Entwurfes von SOA beitragen. Aus diesem Grund soll im Rahmen dieser Dissertation eine Methode zur Anwendung von AD im Entwurf von SOA entwickelt werden und überprüft werden, inwiefern die Vorteile von AD tatsächlich helfen können, ausgewählte Architekturziele zu erreichen.

² Termine der Expertengespräche: Deutsche Post AG: 2006-11-17, BEA Systems GmbH: 2006-09-26, IBM: 2006-09-26, Quimonda AG: 2006-11-22, Sun Microsystems: 2006-11-22 und Software AG: 2006-09-26.

6 Einordnung in die Forschung

Drei Paradigmen charakterisieren die Forschung im Bereich der Informationssysteme und dem Software Engineering – die positivistische, die interpretative und die designorientierte Forschung [GKVi2001, 171 ff.; HMPa2004, 75; VaKu2006].

Das positivistische Paradigma basiert auf der Annahme, dass soziale Welten analog zur Natur aufgefasst und daher nach denselben Prinzipien, die in den Naturwissenschaften verwendet werden, erforscht werden können [GKVi2001, 171]. Erklärungen über Kausalzusammenhänge der Natur werden in Theorien fundiert. Die Vertreter des Positivismus nehmen an, dass die Realität objektiv gegeben ist und über messbare Attribute beschrieben werden kann (empirische Sicht auf die Realität). Diese Attribute sind unabhängig vom Beobachter (Forscher) und seinen Instrumenten. Anliegen der positivistischen Forschung im Bereich der Informationssysteme ist die Entwicklung und die Verbesserung von Theorien, die Erklärungen über Phänomene im Zusammenhang mit Informationssystemen liefern [HMPa2004, 76]. Diese Theorien erklären oder prognostizieren Interaktionen zwischen Mensch, Technologie und Organisation. Ziel der positivistischen Forschung ist die Wahrheitsfindung [HMPa2004, 80].

Das interpretative Paradigma geht von der Annahme aus, dass Wissen gesellschaftlich durch Menschen, die aktiv in den Prozess der Forschung involviert sind, konstruiert wird [GKVi2001, 172]. Forschungsaktivitäten sind das Produkt der Werte des Forschers. Der Zugriff auf die Realität ist nur über gesellschaftliche Konstruktionen, wie Sprache, Bewusstsein oder dem Austausch von Meinungen, möglich. Die philosophische Basis der interpretativen Forschung ist die Hermeneutik und Phänomenologie [Bola1985; GKVi2001, 172]. In der interpretativen Forschung wird versucht, ein Verständnis für die Phänomene der Realität zu erlangen. Dieses Verständnis wird nicht aus der Realität, sondern aus den Meinungen anderer Menschen (Forscher) über diese Phänomene abgeleitet.

Designorientierte Forschung hat ihre Wurzeln in der Konstruktionslehre und den Kunstwissenschaften [Simo1996, 111 ff.]. Es handelt sich dabei im Kern um ein Problemlösungsparadigma, in dessen Rahmen Lösungen kreiert und bewertet werden. Dabei geht es um die Suche nach Innovationen, mit deren Hilfe z. B. die Analyse, das Design, die Implementierung und das Management von Informationssystemen effektiv und effizient durchgeführt werden kann [HMPa2004, 76-77]. Innovationen sind das Ergebnis gestaltungsorientierter Forschung. Sie werden als Artefakt bezeichnet. Das Ziel dieser

Forschungsrichtung ist die Anwendung neuartiger Artefakte zur Lösung bestehender Probleme [HMPa2004, 80].

6.1 Designorientierte Forschung

Design kann sowohl verstanden werden als ein Prozess (d. h. eine Aneinanderreihung von Aktivitäten), als auch als ein Ergebnis (d. h. ein Artefakt) [WWE1992]. Designorientierte Forschung umfasst ein Problemlösungsparadigma dessen Perspektive sich ständig zwischen Designprozess und Artefakt bewegt. Ausgangspunkt designorientierter Forschung ist die Wahrnehmung eines Problems. Der Designprozess ist eine Sequenz von Aktivitäten, die zur Lösung des Problems durchlaufen werden. Am Ende entsteht ein innovatives Ergebnis, ein Artefakt. Innovativ bedeutet, dass dieses Artefakt entweder eine Lösung für ein zuvor ungelöstes Problem repräsentiert oder eine Lösung, die ein bekanntes Problem effektiver oder effizienter löst [HMPa2004, 82]. Aus der anschließenden Evaluation des Artefaktes lassen sich Verbesserungen sowohl für das Artefakt als auch für den Designprozess ableiten. Unter Berücksichtigung dieser Verbesserungen wird der Designprozess anschließend erneut durchlaufen. Dieser „trial-and-error“-Zyklus wird in mehreren evolutionären Schritten durchgespielt, solange bis ein geeignetes Artefakt zur Lösung des Problems gefunden wurde [MMGa2002]. Der Zyklus unterscheidet zwei elementare Schritte, die iterativ durchlaufen werden müssen, dem Designprozess und der anschließenden Evaluation. March und Smith [MaSm1995, 255 ff.] sowie Hevner, March und Park [HMPa2004, 78 ff.] nennen diese Schritte „build“ und „evaluate“. Simon [Simo1996, 128] nennt diese Schritte „generate“ und „test“.

March und Smith [MaSm1995, 255 ff.] unterscheiden die vier Artefakte: Konstrukt, Modell, Methode und Instanz. Konstrukte repräsentieren die Sprache, in welcher Probleme und Lösungen definiert und kommuniziert werden. Modelle verwenden Konstrukte, um eine Situation der realen Welt abzubilden – das Problem und seinen Lösungsraum. Methoden repräsentieren einen Prozess, der beschreibt, wie der Problemlösungsraum durchsucht werden muss, damit ein bestimmtes Problem gelöst werden kann. Methoden können z. B. mathematische Algorithmen, textuelle Beschreibungen von „best practices“ oder Methoden zur Entwicklung von Systemen sein [MaSm1995, 257]. Instanzen sind Konstrukte, Modelle oder Methoden die in einer Umgebung angewendet wurden. Eine Umgebung besteht aus Menschen, Organisationen und der existierenden oder geplanten Technologie [HMPa2004, 79]. Sie definiert den Problemraum, der Ausgangspunkt für die gestaltungsorientierte Forschung ist [Simo1996]. Sie ist somit Ausgangs- und Endpunkt gestaltungsorientierter Forschung. Eine

wichtige Grundlage zur Kreierung und Bewertung neuer Artefakte ist eine Wissensbasis, die Wissen (z. B. Methoden, Theorien, Forschungsframeworks, Modelle und Metriken) enthält, das entweder aus Erkenntnissen früherer Forschung oder aus Referenzdisziplinen stammt [HMPa2004, 80]. Erkenntnisse, die aus einem gestaltungsorientierten Forschungsprojekt gewonnen werden, fließen selbst in die Wissensbasis ein und erweitern diese.

6.2 Einordnung dieser Arbeit

Im Rahmen dieser Arbeit soll eine Methode zur Verbesserung des Entwurfs von SOA erstellt und bewertet werden. Dies ist eine Aufgabe designorientierter Forschung [MaSm1995, 257]. Die Methode ist das Artefakt, das im Sinne der designorientierten Forschung kreiert und evaluiert werden muss. Sie entspricht dem Artefakttyp „Methode“, der von March und Smith [MaSm1995, 257] definiert wurde. Bei der Erstellung dieser Methode wird auf die Wissensbasis AD und Vorgehensmodelle zur Entwicklung von SOA zurückgegriffen. Im Rahmen einer Fallstudie ist geplant, die Anwendbarkeit der zu erstellenden Methode und ihre Ergebnisse zu überprüfen (vgl. Abschnitt 7). Die Methode in der Anwendung entspricht dem Artefakttyp „Instanz“ nach March und Smith [MaSm1995, 257]. Die Umgebung, in welcher die Methode angewendet wird, ist ein reales Entwicklungsprojekt im Entwurf SOA.

7 Forschungsframework und Methodik

Grundlage dieser Arbeit ist ein designorientierter Forschungsansatz basierend auf Takeda et al. [TVTY1990], der so genannte Design Cycle. Dieser Ansatz differenziert fünf Phasen. (Abb. 2).

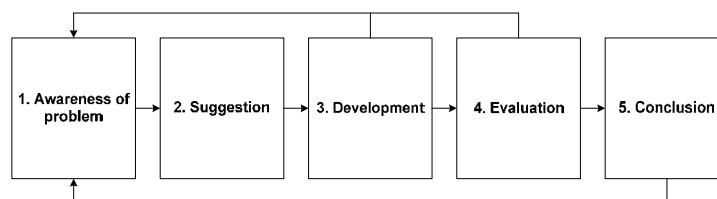


Abb. 2: Design Cycle (in Anlehnung an [TVTY1990, 45])

Die Forschung beginnt mit der Wahrnehmung eines Problems (Awareness of Problem). Ergebnis dieser Phase ist die formale oder informale Beschreibung des Problems und die Begründung eines Forschungsbedarfs, z. B. in Form einer oder mehrerer Forschungsfragen [TVTY1990]. Diese Phase wurde im Rahmen dieser Arbeit bereits durchlaufen. Die Ergebnisse

enthält das hier vorliegende Proposal (vgl. insbesondere Abschnitt 1.1 und 5.1). Das identifizierte Problem besteht darin, dass Vorgehensmodelle zur Entwicklung von SOA die Erreichung bestimmter Architekturziele von SOA nicht ausreichend berücksichtigen. Der abgeleitete Forschungsbedarf spiegelt sich in der Zielstellung (vgl. Abschnitt 1.2) wieder. Im Kern geht es darum, einen Weg zu finden, die Erreichung der Architekturziele von SOA zu verbessern. Ein weiterer Arbeitsschritt dieser Phase ist eine umfassende Ermittlung und Beschreibung der Architekturziele von SOA und die Ermittlung der Abhängigkeiten zwischen diesen Zielen. Dieser Schritt ist notwendig, um später eine begründete Auswahl relevanter Architekturziele vornehmen zu können.

In Phase zwei (Suggestion) wird ein Konzept zur Lösung des Problems erarbeitet. Dieses Konzept kann auf Basis einer vorhandenen Theorie- oder Wissensbasis erstellt werden [HMPa2004, 80; TVTY1990; VaKu2006]. Das Konzept, das im Rahmen dieser Arbeit entwickelt wird, muss eine Methode beschreiben, in deren Rahmen AD für den Entwurf von SOA eingesetzt werden kann, um die Erreichung der Architekturziele von SOA zu verbessern (vgl. Abschnitt 1.2). Die Wissensbasis, auf welche dabei zugegriffen wird, umfasst AD und Vorgehensmodelle zur Entwicklung von SOA. Der erste Schritt dieser Phase umfasst die Ermittlung und Beschreibung von Vorgehensmodellen SOA. Diese Modelle werden miteinander verglichen. Aus den identifizierten Gemeinsamkeiten wird ein idealtypisches Vorgehensmodell zur Entwicklung SOA extrahiert. Anschließend wird das so genannte V-Modell von AD in die entwurfsrelevanten Schritten des idealtypischen Vorgehensmodells SOA integriert [DoSu2000, 279 ff.]. Das V-Modell von AD darf nicht mit Vorgehensmodellen der Softwareentwicklung, zum Beispiel dem V-Modell[®] XT der Bundesbehörden, verwechselt werden. Das V-Modell des AD beschreibt ein grobes Vorgehen, in dessen Rahmen die Konzepte des Axiomatic Design mit jedem beliebigen Vorgehen in der Softwareentwicklung kombiniert werden können [Suh2001, 266]. An einem Fallbeispiel wird die Nutzung der Prinzipien und Hilfsmittel des AD (z. B. das Unabhängigkeits- und Informationsaxiom, die Gesamteinflussmatrix und das Flussdiagramm) in den Entwurfsschritten des Vorgehensmodells SOA demonstriert. Abschließend wird diskutiert, inwieweit die Architekturziele SOA durch Nutzung der Methode erreicht werden können. Ziel dieser Diskussion ist die Selektion relevanter Architekturziele, deren Erreichung durch Nutzung der Methode beeinflusst werden kann.

In der dritten Phase (Development) wird das Lösungskonzept einer Realisierung zugeführt [TVTY1990]. Dabei geht es im Wesentlichen um die Beantwortung der Frage, ob das realisierte Konzept tatsächlich funktionsfähig ist [MaSm1995, 258]. Im Rahmen dieser Arbeit geht es darum, die Anwendbarkeit der auf AD basierenden Methode nachzuweisen [VaKu2006]. Hierzu wird zunächst eine geeignete Evaluationsmethode ausgewählt. Es ist geplant, die Fallstudientechnik zu verwenden [ZeWa1998, 25]. Die auf AD basierende Methode wird hierzu in einem realen Entwicklungsprojekt im Entwurf SOA eingesetzt. Die resultierende Entwurfsspezifikation wird mit Hilfe der Unified Modeling Language (UML) dokumentiert.

Phase vier (Evaluation) umfasst die Bewertung des realisierten Lösungskonzeptes [TVTY1990]. Ziel dieser Phase ist es, herauszufinden, wie gut das realisierte Konzept funktioniert [MaSm1995, 258]. Dabei wird auf Bewertungskriterien zurückgegriffen, die entweder implizit oder explizit in Phase eins (Awareness of Problem) festgelegt wurden. Im Rahmen dieser Arbeit geht es darum, zu prüfen, inwiefern die Architekturziele SOA nach Anwendung der auf AD basierenden Methode tatsächlich erreicht werden können. Hierzu wird zunächst eine geeignete Architekturbewertungsmethode ausgewählt. Qualitative Bewertungsmethoden können zur Bewertung jeglicher Art von Faktoren verwendet werden. Die Architecture Trade-off Analysis Method (ATAM) ist eine strukturierte szenariobasierte Methode zur qualitativen Bewertung von Softwarearchitekturentwürfen [KKBL1998]. Sie berücksichtigt die gegenseitige Beeinflussung zwischen einzelnen Faktoren (z. B. im Falle einer Konflikt- oder Komplementärbeziehung). Es ist geplant, ATAM zur Überprüfung der Zielerreichung im Rahmen dieser Arbeit einzusetzen. Der Kern der Methode umfasst den Aufbau eines hierarchisch strukturierten Qualitätsmodells, den so genannten „Utility Tree“. Er wird auf Basis der in Phase eins ermittelten Ziele von SOA erstellt. Die obersten Knoten des „Utility Tree“ repräsentieren die Architekturziele SOA, die Blattknoten repräsentieren zugehörige Szenarien, die einen Einfluss auf die Zielerreichung haben. Zur Überprüfung, inwiefern die Ziele SOA erreicht wurden, wird anschließend im Rahmen von Workshops mit den Projektbeteiligten überprüft, inwieweit diese Szenarien durch die entworfene Architektur erfüllt werden.

Der Design Cycle endet mit Phase fünf (Conclusion). Hier werden die gesammelten Erkenntnisse der Forschung zusammengetragen und dokumentiert [TVTY1990]. Im Rahmen dieser Arbeit wird die Anwendbarkeit der auf AD basierenden Methode in zukünftigen Entwurfsprojekten für SOA diskutiert. In diese Diskussion fließt die Gegenüberstellung der

Vor- und Nachteile ein, die sich bei der Anwendung der Methode im Rahmen der Fallstudie ergeben haben. Es werden daraufhin Empfehlungen zum Einsatz oder zur Unterlassung der Verwendung der entwickelten Methode formuliert.

Erkenntnisse, die in den Phasen drei, vier und fünf gesammelt werden, können sich auf die Problemwahrnehmung auswirken [TVTY1990]. In diesem Fall ist ein Rücksprung in die Phase eins notwendig. Der gesamte Forschungsansatz kann iterativ in mehreren Zyklen durchlaufen werden. Jede erneute Iteration führt zu einer stärkeren Annäherung an die Lösung des in Phase eins entdeckten Problems. Der Design Cycle funktioniert nach dem Trial and Error – Prinzip, d. h., dass für ein Problem eine Lösung erarbeitet wird. Die Überprüfung des Problemlösungsbeitrages durch Anwendung der Lösung zeigt anschließend, ob diese noch einmal überarbeitet und verbessert werden muss. Auf Grund des begrenzten Zeitrahmens kann in dieser Arbeit nur ein Gesamtdurchlauf des Design Cycle bearbeitet werden. Dies schließt allerdings kleine Iterationen innerhalb des Frameworks nicht aus. Zum Beispiel werden die Architekturziele SOA in der Phase eins ermittelt. Erst nach Durchlauf der Phase zwei wird eine Auswahl relevanter Ziele vorgenommen. Diese Eingrenzung entspricht einem Rücksprung in die Phase eins und einer Verfeinerung der Problemstellung der Arbeit. Sollten Folgeiterationen des Design Cycle notwendig sein, wird im Ausblick dieser Arbeit auf offene Fragen und zukünftige Forschungstätigkeiten hingewiesen.

8 Aufbau

Die Reinschrift dieser Dissertation wird in fünf Kapitel gegliedert. Im ersten Kapitel wird die Einführung in die Themenstellung beschrieben. Der Stand der Forschung wird präsentiert. Davon ausgehend erfolgt die Begründung der Themenstellung. Abschließend werden das Forschungsframework und die methodischen Schritte zur Erreichung der Zielstellung sowie der Aufbau der Arbeit beschrieben. Kapitel zwei und drei sind wichtige Grundlagenkapitel. Sie repräsentieren die Wissensbasis dieser Dissertation. Kapitel zwei umfasst die Beschreibung wichtiger Grundlagen zum Verständnis von SOA für den Kontext dieser Arbeit. Außerdem werden der Entwurf und die Architekturziele von SOA beschrieben. Im Kapitel drei werden die Konzepte und Hilfsmittel des Axiomatic Design aufgeführt. Kapitel vier ist das Kernkapitel dieser Arbeit. Hier wird die Entwicklung der Methode zur Verbesserung der Entwurfsqualität von SOA auf Basis von Axiomatic Design beschrieben. Anschließend wird die Anwendung dieser Methode in einem realen Entwicklungsprojekt im Entwurf SOA dargelegt. Außerdem

wird die Evaluation der Entwurfsergebnisse im Hinblick auf die Erreichung der Architekturziele von SOA mit Hilfe von ATAM beschrieben. In einer abschließenden Bewertung erfolgt die Gegenüberstellung der Vor- und Nachteile, die mit der Anwendung von AD im Entwurf verbunden sind. Im Kapitel fünf werden die Erkenntnisse der Forschung zusammengefasst. Es werden Schlussfolgerungen dargelegt, ob die entwickelte Methode auf Basis von Axiomatic Design tatsächlich für die Verwendung in Entwurfsprojekten von SOA geeignet ist. In einem Ausblick werden offene Forschungsfragen und Hinweise auf zukünftige Forschungstätigkeiten beschrieben.

Literaturverzeichnis

- [Balz1998] Balzert, H.: Lehrbuch der Software-Technik: Software-Management, Softwarequalitätssicherung, Unternehmensmodellierung. Spektrum, Heidelberg et al. 1998.
- [Berl2006] o. V.: SOA in der Praxis: Wie Unternehmen SOA erfolgreich einsetzen. Berlecon Research, Berlin 2006.
- [BHMS2005] Berbner, R.; Heckmann, O.; Mauthe, A.; Steinmetz, R.: Eine Dienstgüte unterstützende Web-Service-Architektur für flexible Geschäftsprozesse. In: Wirtschaftsinformatik 47 (2005) 5, S. 268-277.
- [BiLi2006] Bichler, M.; Lin, K.-J.: Service-Oriented Computing. In: IEEE Computer 39 (2006) 3, S. 99-103.
- [Bola1985] Bolland, R.: Phenomenology: A Preferred Approach to Research in Information Systems. In: Mumford, E. et al. (Hrsg.): Research Methods in Information Systems. Elsevier, Amsterdam 1985, S. 193-201.
- [BuGa2005] Buchmann, I.; Gamber, M.: Methoden zur Unterstützung der Entwicklung einer SOA. In: Cremers, A. B. et al. (Hrsg.): Informatik 2005: Informatik live! GI, Bonn 2005, S. 601-605.
- [CIHi2000] Clapsis, P. J.; Hintersteiner, J. D.: Enhancing Object-Oriented Software Development through Axiomatic Design. In: First International Conference on Axiomatic Design (ICAD2000). Cambridge 2000, S. 272-277.
- [CoRe1996] Cochran, D. S.; Reynal, V. A.: Axiomatic Design of Manufacturing Systems. Massachusetts Institute of Technology, Massachusetts 1996.
- [DeSu2004] Deo, H. V.; Suh, N. P.: Axiomatic Design of customizable automotive Suspension. In: Proceedings of ICAD2004. Seoul 2004.
- [Dijk1982] Dijkstra, E. W.: On the role of scientific thought. In: Dijkstra, E. W. (Hrsg.): Selected writings on computing: a personal perspective. Springer, Berlin, et al. 1982, S. 60-66.
- [DJMZ2005] Dostal, W.; Jeckel, M.; Melzer, I.; Zengler, B.: Service-Orientierte Architekturen mit Web Services: Konzepte - Standards - Praxis. Spektrum, München 2005.
- [DoPa2001] Do, S.-H.; Park, G.-J.: Application of Design Axioms for Glass Bulb Design and Software Development for Design Automation. In: Journal of Mechanical Design 123 (2001) 3, S. 322-329.

- [DoSu1999] Do, S.-H.; Suh, N. P.: Systematic OO Programming with Axiomatic Design. In: IEEE Computer 32 (1999) 10, S. 121-124.
- [DoSu2000] Do, S.-H.; Suh, N. P.: Object-Oriented Software Design with Axiomatic Design. In: Proceedings of ICAD2000. Cambridge 2000, S. 278-284.
- [EAAC2004] Endrei, M.; Ang, J.; Arsanjani, A.; Chua, S. et al.: Patterns: Service-Oriented Architecture and Web Services. <http://www.redbooks.ibm.com>, 2004, Abruf am 2006-06-27.
- [EKAP2005] Erradi, A.; Kulkarni, N.; Anand, S.; Padmanabhuni, S.: Designing Reusable Services: An Experimental Perspective for the Securities Trading Domain. In: Chung, J.-Y. et al. (Hrsg.): Proceedings of the First International Workshop on Design of Service-Oriented Applications (WDSOA'05). IBM Research Division, Amsterdam 2005, S. 25-32.
- [EnNo2000] Engelhardt, F.; Nordlund, M.: Strategic Planning based on Axiomatic Design. In: Proceedings of ICAD2000. Cambridge 2000, S. 26-34.
- [Erl2005] Erl, T.: Service-Oriented Architecture: Concepts, Technology, and Design. Prentice Hall PTR, Upper Saddle River et al. 2005.
- [FiSt2007] Fiege, R.; Stelzer, D.: Analyse des Beitrages von Axiomatic Design zum Entwurf Serviceorientierter Architekturen. In: 8. Internationale Tagung Wirtschaftsinformatik. Karlsruhe 2007.
- [Gall2004] Gallas, E.: Der Aufbau eines Service Life Cycle Managements für eine Service Orientierte Architektur als Brücke zwischen Geschäftsprozesses und IT Integration. In: Aier, S. et al. (Hrsg.): Enterprise Application Integration - Serviceorientierung und nachhaltige Architekturen. GITO, Berlin 2004, S. 231-278.
- [GKVi2001] Gregg, D.; Kulkarni, U.; Vinze, A.: Understanding the Philosophical Underpinnings of Software Engineering Research in Information Systems. In: Information Systems Frontiers 3 (2001) 2, S. 169-183.
- [HaSc2006] Hagen, C.; Schwinn, A.: Measured Integration - Metriken für die Integrationsarchitektur. In: Schelp, J. et al. (Hrsg.): Integrationsmanagement. Planung, Bewertung und Steuerung von Applikationslandschaften. Springer, Berlin et al. 2006, S. 267-292.
- [HMPa2004] Hevner, A. R.; March, S. T.; Park, J.: Design Science in Information Systems Research. In: MIS Quarterly 28 (2004) 1, S. 75-105.
- [Jams2004] Jamshidnezhad, B.: Towards a Rational Basis for User Interface Design Methods. In: Proceedings of ICAD2004. Seoul 2004.
- [KKBL1998] Kazman, R.; Klein, M.; Barbacci, M.; Longstaff, T. et al.: The Architecture Tradeoff Analysis Method. In: Proceedings of the Fourth IEEE International Conference on Engineering of Complex Computer Systems (ICECCS). Monterey 1998, S. 68-78.
- [KoHB2005] Kotonya, G.; Hutchinson, J.; Bloin, B.: A Method for Formulating and Architecting Component- and Service-Oriented Systems. In: Stojanovic, Z. et al. (Hrsg.): Service-Oriented Software System Engineering: Challenges and Practices. IGP, Hershey et al. 2005, S. 155-181.
- [LaMB2005] Laures, G.; Meyer, H.; Breest, M.: An Engineering Method for Semantic Service Applications. In: Chung, J.-Y. et al. (Hrsg.): Proceedings of the First International Workshop on Design of Service-Oriented Applications (WDSOA'05). IBM Research Division, Amsterdam 2005, S. 79-86.

- [MaBe2006] Marks, E. A.; Bell, M.: Service-Oriented Architecture: A Planning and Implementation Guide for Business and Technology. Wiley, Hoboken et al. 2006.
- [MaSm1995] March, S. T.; Smith, G. F.: Design and natural science research on information technology. In: Decision support systems 15 (1995) 4, S. 251-266.
- [Mert2006] Mertens, P. (Hrsg.): Methoden und Nachhaltigkeit in der Wirtschaftsinformatik. Arbeitbericht Nr. 1/2006, Universität Erlangen-Nürnberg 2006.
- [MMGa2002] Markus, M. L.; Marjchrzak, A.; Gasser, L.: A Design Theory for Systems that Support Emergent Knowledge Processes. In: MIS Quarterly 26 (2002) 3, S. 179-212.
- [PaYa2002] Papazoglou, M. P.; Yang, J.: Design Methodology for Web Services and Business Processes. In: Lecture notes in computer science 2444 (2002), S. 54-64.
- [SiHu2005] Singh, M. P.; Huhns, M. N.: Service-Oriented Computing: Semantics, Processes, Agents. Wiley, Chichester et al. 2005.
- [Simo1996] Simon, H. A.: The sciences of the artificial. 3. Aufl., MIT Press, Cambridge, Mass. et al. 1996.
- [SSLD2005] Steen, M. W. A.; Strating, P.; Lankhorst, M. M.; Doest, H. W. L.: Service-Oriented Enterprise Architecture. In: Stojanovic, Z. et al. (Hrsg.): Service-Oriented Software System Engineering: Challenges and Practices. IGP, Hershey et al. 2005, S. 132-153.
- [StDu2000] Steward, D.; Tate, D.: Integration of Axiomatic Design and Project Planning. In: Proceedings of ICAD2000. Cambridge 2000, S. 285-289.
- [SuDo2000] Suh, N. P.; Do, S.-H.: Axiomatic Design of Software Systems. In: Annals of the CIRP 49 (2000) 1, S. 95.
- [Suh2001] Suh, N. P.: Axiomatic Design: Advances and Applications. Oxford, New York 2001.
- [Suh1998] Suh, N. P.: Axiomatic Design Theory for Systems. In: Research in Engineering Design 10 (1998) 4, S. 189-209.
- [Suh1990] Suh, N. P.: The Principles of Design. Oxford, New York et al. 1990.
- [TVTY1990] Takeda, H.; Veerkamp, P.; Tomiyama, T.; Yoshikawam, H.: Modeling Design Processes. In: AI Magazine 11 (1990) 4, S. 37-48.
- [VaKu2006] Vaishnavi, V.; Kuechler, B.: Design Research in Information Systems. <http://www.isworld.org/Researchdesign/drisISworld.htm>, 2006, Abruf am 2006-11-07.
- [WKAd2000] Werneman, A.; Kjellberg, A.; Adman, M.: Application of Axiomatic Design in operational Development. In: Proceedings of ICAD2000. Cambridge 2000.
- [WWEI1992] Walls, J. G.; Widemeyer, G. R.; El Sawy, O. A.: Building an Information System Design Theory for Vigilant EIS. In: Information systems research 3 (1992) 1, S. 36-59.
- [YiPa2004] Yi, J.-W.; Park, G.-J.: Software Development of a Sequential Algorithm with Orthogonal Arrays (SOA) using Axiomatic Design. In: Proceedings of ICAD2004. Seoul 2004.
- [ZeWa1998] Zelkowitz, M. V.; Wallace, D. R.: Experimental Models for Validating Technology. In: IEEE Computer 31 (1998) 5, S. 23-31.
- [ZSWP2005] Zimmermann, O.; Schlimm, N.; Waller, G.; Pestel, M.: Analysis and Design Techniques for Service-Oriented Development and Integration. In: Cremers, A. B. et al. (Hrsg.): Informatik 2005: Informatik live! GI, Bonn 2005, S. 606-611.

Modellierung und Messung des Einflusses von RFID-Technologie auf die Produktivität

Daniel Gille

Institut für Informatik und Gesellschaft – Abteilung Telematik
Albert-Ludwigs-Universität Freiburg i. Brsg.
daniel.gille@iig.uni-freiburg.de

1 Einführung

RFID-Technologie (*Radio Frequency Identification*) ermöglicht durch automatisierte Informationserfassung und verbesserte Informationen über physikalische Objekte deutliche Produktivitätssteigerungen, die in der Praxis bereits ansatzweise realisiert werden [Ange2005]. Methoden zur systematischen Messung der verschiedenen Arten einer RFID-bedingten Produktivitätssteigerung wurden bislang nicht entwickelt. Eine aktuelle Studie des Marktforschungsunternehmens Aberdeen Research kommt zu dem Ergebnis, dass gerade die ungenügende Messbarkeit des Nutzens der RFID-Technologie ein wesentliches Investitionshindernis für viele Unternehmen darstellt [Aber2006]. Hinsichtlich der Produktivitätsmessungen beim Einsatz von Informationstechnologie (*IT*) in der Vergangenheit muss konstatiert werden, dass den beiden wesentlichen Funktionen der betrieblichen Leistungsmessung, der Gewinnung von erfolgsrelevanten Informationen und Steuerung von Geschäftsprozessen, nur unzureichend Rechnung getragen wurde. Die Fokussierung auf Firm-Level-Analysen und Multifaktorproduktivität als alleinige Kennzahl für den Einfluss der IT hat zur Folge, dass Informationen über die optimale Verwendungsart der Technologie und Möglichkeiten der Steuerung in nur ungenügendem Umfang gewonnen werden können. Dies ist darauf zurückzuführen, dass Firm-Level-Analysen von den Geschäftsprozessen abstrahieren, in denen durch den IT-Einsatz ökonomischer Wert generiert wird. Prozesse stellen aus Sicht unternehmensweiter Produktivitätskennzahlen vielmehr eine „Black Box“ dar. Aus diesem Grund wird in dem Dissertationsvorhaben eine *prozessorientierte* Analyse der Produktivitätssteigerungen beim RFID-Einsatz angestrebt. Auf Grundlage einer erweiterten Definition des Produktivitätsbegriffs (als „Performance“ bzw. „Leistung“) sollen auf Prozessebene finanzielle und nicht-finanzielle Kennzahlen identifiziert und in einem multidimensionalen Messsystem zusammengeführt werden. Das zu entwickelnde Messsystem soll

verschiedene Arten einer RFID-bedingten Leistungssteigerung erfassen und im Praxisbetrieb Erkenntnisse hinsichtlich der optimalen Prozessintegration liefern.

2 Ausgangslage: Leistungssteigerung und -Messung beim RFID-Einsatz

2.1 Leistungssteigerung beim RFID-Einsatz

2.1.1 Potenziale zur Leistungssteigerung

RFID-Technologie stellt aufgrund ihrer technischen Funktionalität und der steigenden Verbreitung in zahlreichen betrieblichen Einsatzgebieten eine erste relevante Ausprägung des „allgegenwärtigen Rechnens“ (*Ubiquitous Computing; UC*) dar. Die Technologie erlaubt das sichtkontaktlose und automatisierte Erfassen und Verarbeiten von objektbezogenen Informationen in Echtzeit. Die Informatisierung der physikalischen Welt und all ihrer Gegenstände wird möglich [Muel2003] [Muel2006]. Die einem Unternehmen zur Entscheidungsfindung verfügbaren Informationen werden kostengünstiger verfügbar und zudem qualitativ hochwertiger (geringere Fehlerhäufigkeit; sofortige Verfügbarkeit; Objektivität; höhere Datengranularität [Tell2006]). In letzter Instanz wird es möglich, dass Informationssysteme über ein stets aktuelles und hinsichtlich betriebswirtschaftlicher Entscheidungsgrößen vollständiges Realitätsmodell der Umgebung verfügen („*permanente Vollerfassung*“). In vielen Bereichen kann somit erstmals auf der Grundlage von *tatsächlichen* (anstelle von buchhalterisch fortgeschriebenen) Bestands- und Bewegungsdaten von an den Leistungserstellungsprozessen beteiligten Objekten eines Unternehmens geplant, produziert und auf Veränderungen reagiert werden [FlCD2005]. Eine produktivere Gestaltung von Unternehmensprozessen wird über drei verschiedene, sich nicht gegenseitig ausschließende Effekte möglich [MoGK1996] [Tell2006].

- *Automatisierungseffekt*: Vormalig manuelle Prozesse der Informationserfassung und –weiterleitung können mit RFID automatisiert werden. Bspw. kann das Erfassen einer Palettennummer am Wareneingang durch einen Mitarbeiter mit einem mobilen Barcodescanner entfallen.
- *Informationseffekt*: Eine Erhebung zusätzlicher bzw. qualitativ besserer Informationen durch den RFID-Einsatz ermöglicht eine Verbesserung von Koordinations-

entscheidungen. So können etwa in der Lagerhaltung automatisierte Nachbestellungen auf Grundlage einer kontinuierlichen Objekterfassung durchgeführt werden. Unproduktive Tätigkeiten, bspw. das vergebliche Aufsuchen einer Entnahmestelle bei nicht vorrätigen Produkten, können vermieden werden.

- *Transformationseffekt*: Aufgrund dieser verbesserten Informationsgrundlage können bestehende Unternehmensprozesse umgestaltet werden. So kann bspw. in Distributionszentren des Einzelhandels die herkömmliche Kommissionierung teilweise durch die Bereitstellung und Umverteilung filialgerecht vorkommissionierter Paletten an der Verladerampe ersetzt werden (*Cross-Docking* [GiHa2006]).

2.1.2 Praktische Umsetzung der Potenziale

In der betrieblichen Praxis sind bereits erste Ansätze der Umsetzung dieser Produktivitätseffekte erkennbar. Zu diesem Ergebnis kommt die empirische Studie *RFID in Deutschland*¹ [StGi2006]. Allerdings wird bisher erst die Automatisierung der Informationserfassung in breiterem Ausmaß realisiert. So geben 22 % der befragten Unternehmen an, diesen Effekt bspw. in Form einer automatisierten Bestandserfassung zu nutzen. Eine Verbesserung der Koordination (*Informationseffekt*), z.B. in Form einer Nutzung von RFID zur Bevorratung und Bedarfsprognose, wird von einem geringeren Anteil der Unternehmen realisiert.²

Insgesamt konnte ein Anteil von 45 % der befragten Unternehmen identifiziert werden, die bereits RFID-Technologie einsetzen (weitere 46 % planen dies für die kommenden zwei Jahre bzw. sehen den Einsatz als grundsätzlich geeignet an). Als Unternehmensbereich, für den RFID-Technologie die größte Relevanz besitzt, konnte die *Logistik und Lagerhaltung* und darin als Anwendung die *Kommissionierung* identifiziert werden.

Unter den wichtigsten Zielen eines RFID-Einsatzes in der Logistik können mit der Verbilligung der Inventur, Minimierung von Beständen und Durchlaufzeiten sowie der Reduzierung der Personalkosten explizit *Produktivitätssteigerungen* identifiziert werden [StGi2006]. Auch andere Studien unterstreichen die Bedeutung produktivitätsbezogener Zielsetzungen [BoDL2004].

¹ Die Studie „RFID in Deutschland“ wurde als Zufallserhebung für Branchen konzipiert, in denen RFID derzeit besonders relevant ist. Die vorgestellten Ergebnisse basieren auf den Angaben von 68 Unternehmen, die vornehmlich in den Bereichen Handel, Verarbeitendes Gewerbe und Dienstleistungen (z.B. Logistik) angesiedelt sind.

² Der Transformationseffekt war nicht Gegenstand besagter Studie.

2.2 Leistungsmessung beim Einsatz von IT und RFID

2.2.1 Probleme bei der Performance-Messung von IT

Die Messung der ökonomischen Auswirkungen von IT steht im Mittelpunkt einer Vielzahl von größtenteils empirischen Untersuchungen. Obwohl IT ihre Wirkung primär auf der Ebene von Geschäftsprozessen entfaltet, wurden bislang in erster Linie Versuche unternommen, die Auswirkungen der Technologie durch den aggregierten Vergleich von Kennzahlen auf Firmen-, Branchen- und Länderebene nachzuweisen [DeGK2003]. Zwar kann bspw. durch den Vergleich von IT-Investitionen und Multifaktorproduktivitäten einzelner Firmen [BrHi1998] systematisch nachgewiesen werden, dass IT über die betrachteten Unternehmen hinweg im Durchschnitt zu einer Produktivitätssteigerung beiträgt. *Wie* diese Steigerung in einzelnen Unternehmen erreicht wurde, kann aufgrund der Abstraktion von den zugrunde liegenden Geschäftsprozessen jedoch nicht ermittelt werden [MoGK1996]. Es können lediglich auf abstrakter Ebene komplementäre organisatorische Veränderungen identifiziert werden, welche Unternehmen mit hoher Multifaktorproduktivität im Gegensatz zu ihren weniger „erfolgreichen“ Mitbewerbern (mit vergleichbarem IT-Einsatz) durchgeführt haben (z.B. Dezentralisierung; neue Anreizsysteme [BrHi1998]). Handlungsempfehlungen zur praktischen Umsetzung von IT-Potenzialen in Unternehmensprozessen können hieraus nur eingeschränkt gewonnen werden [MoGK1996], zumal beobachtbar ist, dass einzelne Firmen aus IT-induzierten Produktivitätsfortschritten keine langfristigen Wettbewerbsvorteile ziehen können [Carr2004]. Ähnliches ist für die RFID-Technologie zu erwarten: Als essenzieller Bestandteil der Erbringung logistischer Leistungen wird sie langfristig auf Firmenebene zu keiner Verbesserung der strategischen Position beitragen, sondern durch Leistungsverbesserungen auf Prozessebene die Konkurrenzfähigkeit sichern. Eine verstärkte Konzentration auf Prozesse ist aus diesem Grunde erstrebenswert.

2.2.2 Bisherige Ansätze der Performance-Messung beim RFID-Einsatz

Analysen von RFID-Anwendungen konzentrieren sich bislang auf Einzelfallbeispiele, die die unstrukturierte Darstellung individueller Anwendungen der RFID-Technologie, vornehmlich in der Logistik, zum Gegenstand haben (z.B. [Info2006]). Diese Fallstudien zeigen für ausgesuchte Einzelfälle Einsatzmöglichkeiten der RFID-Technologie auf. Vereinzelt wird der Versuch einer Messung des Nutzens der Technologie (z.B. in Form des ROI) unter stark vereinfachenden Annahmen vorgenommen (z.B. in [GiHa2006] [Stra2005]).

Erste systematische Ansätze zur Quantifizierung der in 2.1.1 beschriebenen Produktivitätseffekte eines RFID-Einsatzes bietet [Tell2006], der zur Messung des Nutzens einer Automatisierung von Informationserfassung (Automatisierungseffekt) im Rahmen eines fiktionalen Beispiels die Prozesskostenrechnung heranzieht. Eine nahezu identische Vorgehensweise wählen [Subi2003] und [Laub2006] zur Berechnung des ROI einer realen RFID-Anwendung im Wareneingang eines Distributionszentrums. Darüber hinausgehende Ansätze zur Quantifizierung des Informations- und Transformationseffekts beschränken sich auf die analytische Ermittlung theoretischer Nutzenpotenziale, die sich aus einer Reduktion von Ungenauigkeiten in den Bestandsdaten (Informationseffekt) sowie einer Prozessneugestaltung zur automatisierten Nachbefüllung von Regalen im Einzelhandel (Transformationseffekt) ergeben [Tell2006]. Ansätze zur Messung dieser Effekte in realen Logistikprozessen werden nicht aufgezeigt. Darüber hinaus sind die angewendeten Modelle auf singuläre Zielgrößen und Anwendungsszenarien ausgerichtet (Personaleinsparungen im Wareneingang [Subi2003] [Laub2006]; Sicherheitspuffer in der Lagerhaltung [Tell2006]; optimaler Zeitpunkt einer Regalnachbefüllung im Einzelhandel [Tell2006]), was ihre Eignung für eine umfassende Leistungsmessung und -bewertung von Logistikprozessen stark eingeschränkt. Weitere branchen- und anwendungsspezifische Analysen fokussieren in ähnlichem Ausmaß auf spezifische Zielgrößen [Ditt2006] [Stra2005].

Prozessübergreifend anwendbare Metriken zur Leistungsbewertung von Logistikprozessen beim RFID-Einsatz existieren bislang nicht. Dies erschwert eine adäquate Bewertung des Nutzens der Technologie und kann aufgrund fehlender Erkenntnisse hinsichtlich der Integration in die Prozessgestaltung und Gesamtstrategie eine Einführung verhindern.

3 Zielsetzung der Arbeit

Ziel des Dissertationsvorhabens ist, ein Messsystem zur systematischen Erfassung des RFID-Einflusses auf die Leistungsfähigkeit von Logistikprozessen zu entwickeln. Ausgehend von den aufgezeigten Problemen bisheriger Performance-Messungen von IT soll das zu entwickelnde Messsystem folgende Attribute aufweisen:

Prozessbasiert: Das Messsystem soll alternativ zu bisherigen Ansätzen den Beitrag der RFID-Technologie zur Leistungssteigerung nicht durch die Erfassung der Firmen-Performance, sondern durch Kennzahlen auf der Ebene von *Prozessen* und deren *Komponenten* erfassen.

Multidimensional: Sowohl finanzielle als auch nicht-finanzielle Kennzahlen (Kosten, Qualität, Flexibilität etc.) sollen berücksichtigt werden und eine Anpassung an verschiedene strategische und operationelle Zielsetzungen ermöglichen.

Anwendungsübergreifend: Die Modellierung abstrakter Prozesse mit generischen Aktivitäten, die im Zuge der Anpassung an konkrete Anwendungen spezifiziert werden, dient als Ausgangspunkt für die anwendungsübergreifende Performance-Messung. Aufgrund des bisherigen Status der RFID-Nutzung werden Logistikanwendungen betrachtet [StGi2006], die über eine ausreichende Anwendungsreife verfügen und zudem große Potenziale der Leistungssteigerung aufweisen. Dies trifft auf die Teilprozesse „Wareneingang“, „Warenausgang“ und insbesondere „Kommissionierung“³ zu. Eine weitere im Rahmen des Dissertationsvorhabens betrachtete Anwendung stellt der Koordinationsmechanismus „Kanban“⁴ dar. Die praktische Anwendung und Evaluierung des Messsystems in realen betrieblichen Anwendungsszenarien soll eine Quantifizierung RFID-induzierter Produktivitätssteigerungen in der Praxis ermöglichen.

4 Bestehende Ansätze

Die Entwicklung des Messsystems baut auf verschiedenen ökonomischen und wirtschaftsinformatischen Ansätzen auf. Zunächst sind die Anforderungen und Zielsetzungen von Ansätzen der Performance-Messung von Relevanz (Abschnitt 4.1). Es kann gezeigt werden, dass gerade die RFID-Technologie eine Vielzahl strategischer Zielsetzungen unterstützt. Dies muss im Design entsprechender Messsysteme berücksichtigt werden. Daran anknüpfend wird zwischen finanziellen (Abschnitt 4.2) und nicht-finanziellen (Abschnitt 4.3) Kennzahlen bzw. Messsystemen unterschieden. Um eine Kombination geeigneter Kennzahlen methodisch ableiten zu können, sind desweiteren Ansätze zur Modellierung von Prozessen (Abschnitt 4.4) sowie die Abbildung

³ In der Kommissionierung wurden bereits erste RFID-Anwendungen realisiert [RoVo2006]. Die auftragsbezogene Zusammenstellung von Teilmengen aus einer Gesamtmenge von Gütern (Sortiment) ist mit einem hohen Koordinationsaufwand verbunden und ohne manuelle Arbeit auf absehbare Zeit nicht realisierbar [Schu1999]. Die qualitative Verbesserung von Informationen durch RFID-Einsatz stellt hier weitergehende Koordinationsverbesserungen in Aussicht. Zudem stellt die Kommissionierung aufgrund des hohen Grades an Standardisierung der Ablauforganisation eine geeignete Basis für Vergleiche über Unternehmensgrenzen hinweg dar.

⁴ Bei Anwendung des Kanban-Prinzips werden in selbststeuernden Regelkreisen zwischen einzelnen Wertschöpfungsstufen (Quellen und Senken) nachfrageorientierte Impulse zur Nachproduktion bzw. Auslieferung von Produkten ausgetauscht (Pull-Prinzip). Auch hier wurden bereits erste RFID-Anwendungen entwickelt [Ibol-2005], welche in erster Linie den Informationsaustausch via Kanban-Karten elektronisch ersetzen und dadurch schneller und fehlerfreier machen.

der Auswirkungen eines RFID-Einsatzes auf Prozesse und deren Komponenten (Abschnitt 4.5) von Bedeutung.

4.1 Ziele von Performance-Messung

Verfahren der betrieblichen Leistungsmessung unter Anwendung finanzieller Kennzahlen werden bereits seit Beginn des 20. Jahrhunderts eingesetzt [Chan1977]. Es besteht ein verbreiteter Konsens darüber, dass diese Metriken nur begrenzt aussagekräftig sind⁵, Fehlverhalten begünstigen [Neel1999] und Unternehmensstrategien unzureichend berücksichtigen [GhNo1996]. Tiefgreifende Veränderungen ökonomischer Rahmenbedingungen seit Beginn der 1980er Jahre, darunter die stetig schwindende Bedeutung menschlicher Arbeit aufgrund von Prozessautomatisierung, intensivierter (internationaler) Wettbewerb, zahlreiche Initiativen zur Prozessverbesserung (Just-In-Time, Lean Production, Total Quality Management, Efficient Consumer Response...), veränderte Anforderungen von Kunden und staatlichen Institutionen sowie der Einsatz von Informationstechnologie, erfordern vielmehr eine *integrierte Messung multidimensionaler Zielgrößen*. Neben der Kostenreduktion wird die Identifikation nicht-finanzieller Differenzierungsmerkmale (Flexibilität, Qualität, Innovation, Lernfähigkeit etc.) zunehmend bedeutend [KaNo1992] [Neel1999]. Zudem sollen Kennzahlen das Verständnis der eigenen Prozesse fördern [Park2000] und aufgrund ihres Einflusses auf das Verhalten von Organisationsmitgliedern verstärkt zur Umsetzung unternehmensspezifischer Strategien eingesetzt werden.⁶ Als Folge dieser Entwicklungen haben multidimensionale, sowohl finanzielle als auch nicht-finanzielle Kennzahlen in einem „balancierten“ Verhältnis berücksichtigende Performance-Messsysteme (*Balanced Scorecard, Performance Measurement Matrix etc.*) eindimensionale und zumeist finanzielle Kennzahlen weitgehend abgelöst [BoFW2003].

Für die Entwicklung eines Systems zur Messung des Einflusses von RFID-Technologie müssen die Erweiterungen und Differenzierungen in den Ansätzen der Performance-Messung berücksichtigt werden, da die Technologie über verschiedene Effekte zu einer Leistungssteigerung beitragen (s. Abschnitt 2.1) und eine Vielzahl von Zielen unterstützen kann. So begünstigt der Automatisierungseffekt eine weitergehende Substitution menschlicher Arbeit durch Kapital (Kostensenkung). Darüber hinaus tragen qualitativ verbesserte Informationen (Informationsef-

⁵ Bspw. werden externe Faktoren wie Kundenbedürfnisse und die Performance von Wettbewerbern nicht berücksichtigt [KaNo1992]. Zudem besteht eine Bewertungsproblematik bei Preisschwankungen [Tang2004].

⁶ So sollen Performance-Maße als „Ersatz“ für mittlere Managementebenen bei verflachten Unternehmenshierarchien die Kommunikation zwischen Unternehmensführung und Angestellten verbessern [Neel1999] und auch zwischen Mitarbeitern einer Hierarchieebene durch Präzisierung der Kommunikation ein gemeinsames Verständnis der zu erreichenden Ziele fördern [SiZa1995].

fekt), bspw. durch den Einsatz von RFID zum Management von Spezialladungsträgern in der Automobilbranche, zu einer Umsetzung strategischer Initiativen der Prozessverbesserung (z.B. Just-in-time) und einer verstärkten Kundenorientierung bei [Stra2005]. Auch die Möglichkeit der Umgestaltung von Prozessen (Transformationseffekt) mit dem Ziel von Prozessverbesserungen und Produktinnovationen steht im Einklang mit neuen Ansätzen der Performance-Messung (Differenzierung von Wettbewerbern, Steigerung der Kundenzufriedenheit etc.).

4.2 Finanzielle Kennzahlen

Traditionelle Leistungskennzahlen geben mehrheitlich in monetarisierten Einheiten Aufschluss über die Performance ökonomischer Einheiten (Prozesse, Unternehmen, Branchen, Staaten etc.). Prinzipiell kann zwischen Produktivitätsmaßen i.e.S. und Profitabilitätsmaßen unterschieden werden [Tang2004].

Produktivitätsmaße: Produktivität im herkömmlichen Sinn ist als Verhältnis von Output zu Input definiert [Tang2002]. Der Output einer wirtschaftlichen Einheit in Relation zu einer spezifischen Input-Quelle (z.B. Arbeit oder Kapital) wird als *Faktorproduktivität* bezeichnet. Die Hauptkritik an der Verwendung dieser Metrik beruht auf die nicht gegebene Berücksichtigung von Interdependenzen zwischen verschiedenen Input-Quellen (z.B. Substitution von Arbeit durch Kapital). So kann in Extremfällen trotz gemessener Produktivitätssteigerung eines Inputfaktors die Stagnation oder Verminderung der Gesamtproduktivität unbemerkt bleiben. Im Gegensatz hierzu stellt die *Multifaktorproduktivität* dem Output die Gesamtheit aller Input-Faktoren gegenüber. Obwohl mit dieser Vorgehensweise eine näherungsweise Bestimmung der Gesamtleistung einer wirtschaftlichen Einheit erreicht werden kann, bestehen zahlreiche Probleme bei der Messung und Bewertung der Input- und Output-Faktoren in der Praxis [Gros1993]. Daraus resultierend sind Multifaktorproduktivitäten oftmals inakkurat und erlauben zudem kaum Rückschlüsse auf die für eine Leistungssteigerung verantwortlichen Faktoren [Tang2004]. Herkömmliche Produktivitätsmaße sind aus diesen Gründen für die weitere Untersuchung nur eingeschränkt relevant.

Profitabilitätsmaße: Profitabilitätsmaße werden in zahlreichen verschiedenen Ausprägungen angewendet. Besonders gebräuchliche Metriken sind der *Return on Investment* (ROI) sowie die *Umsatzrendite* (Return on Sales). Daneben ist der Ansatz des *Activity-Based Costing* (ABC) ein populäres Verfahren zur Identifikation von Kostentreibern innerhalb von Prozessen [Ka-

Co1998]. Die Analyse der mit Prozesskomponenten (Aktivitäten) assoziierten Kosten und deren Zuordnung zu spezifischen Prozessen und Produkten macht dieses Verfahren für die Bewertung der RFID-Technologie besonders attraktiv. Insbesondere der *zeitbasierte ABC-Ansatz* [KaAn2003] findet in der Nutzenbewertung des Automatisierungseffekts der RFID-Technologie bereits Anwendung [Subi2003] [Laub2006] [Tell2006].

4.3 Nicht-finanzielle Kennzahlen(-systeme)

Die Bedeutung nicht-finanzieller Kennzahlen nimmt in jüngeren Forschungsarbeiten zur Performance-Messung stetig zu. Während finanzielle Metriken ursprünglich zur Kontrolle und Reduktion von direkten Arbeitskosten in der Ära der Massenproduktion mit einer geringen Anzahl standardisierter Produkte eingeführt wurden, erlauben nicht-finanzielle Kennzahlen und deren Kombination in konzeptionellen Systemen die Messung im Rahmen zunehmend komplexer Unternehmensstrategien (Lean Production, Mass Customisation etc.). Wichtige Ausprägungen nicht-finanzieller Kennzahlen sind zeitbasierte Produktivitätsmaße sowie die in der jüngeren Vergangenheit entwickelten konzeptionellen Performance-Messsysteme [Tang2004].

Zeitbasierte Produktivitätsmaße: Zur Umgehung des Problems der monetären Bewertung von Performance werden verschiedene zeitbasierte Produktivitätsmaße vorgeschlagen. Ein Maß zur Erfassung der Leistung automatisierter und halbautomatisierter Prozesse ist die *Overall Equipment Effectiveness* (OEE). Ziel ist die Reduktion bzw. Eliminierung von Defekten, Verlangsamungen und Unterbrechungen des Produktionsablaufs aufgrund von Störungen und Anpassungen [ChSh2000]. Die OEE ist definiert als Produkt aus Verfügbarkeit, Leistungseffizienz und Qualität [Naka1988]. Das Gegenstück der OEE zur Bewertung manueller Prozesse ist die *Manual Assembly Efficiency* (MAE) [Pete2000]. Analog zur OEE werden Zeitmaße der Qualität (Nachbearbeitungszeit), Auslastung (Leerlaufzeit) und Verfügbarkeit (ideale Durchlaufzeit, gesamte verfügbare Zeit) in einer Kennzahl zusammengeführt. Zur Analyse des Einflusses von RFID-Technologie in Logistikprozessen ist sowohl die Messung in automatisierten (z.B. Informationserfassung im Wareneingang) als auch in manuellen Prozessen (z.B. Kommissionierung) notwendig.

Konzeptionelle Performance-Messsysteme: Konzeptionelle Messsysteme sollen angesichts der zunehmenden Komplexität betrieblicher Entscheidungssituationen einen umfassenderen Blick auf verschiedene relevante Größen ermöglichen. Dies erfolgt in erster Linie durch die Integrati-

on nicht-finanzieller Kennzahlen (z.B. Durchlaufzeit, Anzahl von Kundenbeschwerden, Häufigkeit von Nachbearbeitungen) [MeSt2000]. Prinzipiell kann zwischen Systemen unterschieden werden, die einen Rahmen für die Kategorisierung, Auswahl und Anordnung von Kennzahlen vorgeben und Methoden, die primär auf die Individualität von Unternehmen abstellen und dementsprechend ein freies unternehmensspezifisches Design von Kennzahlensystemen unterstützen. Zu ersterer Gruppe sind u.a. die *Balanced Scorecard* [KaNo1992] und die *Performance Measurement Matrix* [KeEJ1989] zu zählen. Gemeinsam ist diesen Systemen die Definition bestimmter Klassen bzw. Perspektiven (z.B. intern/extern, kostenorientiert/nicht-kostenorientiert, Kundensicht, Innovationsperspektive etc.), innerhalb derer Ziele definiert und entsprechende Kennzahlen ausgewählt werden. Ein Vorteil dieser differenzierteren Betrachtungsweise besteht darin, dass nachvollzogen werden kann, worauf eine Verbesserung einzelner Kennzahlen zurückzuführen ist [KaNo1992]. Auf der anderen Seite ist als Beispiel für das konzeptionelle Design von Messsystemen der *Auditing and Enhancing Framework* [MeSt2000] zu nennen. In einem sechsstufigen Prozess können aus einer Vielzahl spezifischer Kenngrößen diejenigen identifiziert werden, die gemäß der individuellen Unternehmensstrategie die aussagekräftigste Bewertung des zu analysierenden Unternehmensprozesses ermöglichen. Eine Audit-Phase sowie die periodische Evaluierung des Systems gewährleisten die kontinuierliche Anpassung des Verfahrens. Gerade aufgrund der Verschiedenartigkeit der Performance-Effekte beim RFID-Einsatz sind konzeptionelle Messsysteme in besonderem Maße für deren Erfassung relevant.

4.4 Methoden der Prozessmodellierung

Die Auswahl geeigneter und exakter Kennzahlen für eine Messung des RFID-Einflusses erfordert eine Lokalisierung möglicher Ansatzpunkte einer Performance-Steigerung innerhalb von Prozessen. Es müssen also auf einer Subprozessebene einzelne Prozesselemente auf ihre grundlegende „Verbesserungsfähigkeit“ mit Hilfe der RFID-Technologie überprüft werden. Ein wichtiges Hilfsmittel hierzu stellen Instrumente der Prozessmodellierung zur dar. So kann ein Prozess als eine Anordnung von Arbeitsaktivitäten mit klar definierten In- und Outputs [Dave1993] in die grundlegenden Bestandteile *Aktivitäten*, *Ressourcen* und *Abhängigkeiten* untergliedert werden [Malo1999].

Aktivitäten: Die im Rahmen eines Prozesses durchgeführte Gesamttransformation von Input zu Output kann schrittweise in einzelne Subprozesse und Aktivitäten zerlegt werden, in denen Teil-

le der Gesamtleistung erbracht werden. Im Rahmen der *objektorientierten Geschäftsprozesssimulation* können verschiedene generische Typen von Aktivitäten unterschieden werden, welche für die Modellierung von Fertigungs- und Logistikprozessen relevant sind. Dazu zählen u.a. die Aktivitäten „*Assemble*“ (z.B. Zusammenbau von Einzelteilen), „*Batch*“ (z.B. Erstellung von Chargen) und „*Gate*“ (z.B. Aufstauen von Objekten in einer Warteschlange) [Tuma1996]. Die Modellierung von Aktivitäten ist neben der Simulation von Geschäftsprozessen u.a. Grundlage zur Anwendung des ABC-Ansatzes (s. Abschnitt 4.2).

Abhängigkeiten: Das Verhältnis einzelner Aktivitäten zueinander ist durch die gemeinsame Nutzung von Ressourcen bestimmt und wird anhand von Abhängigkeiten charakterisiert. Es wird zwischen den grundlegenden Abhängigkeitstypen „*Flow*“, „*Sharing*“ und „*Fit*“ unterschieden [Malo1999]. So kann eine Aktivität eine Ressource hervorbringen, welche von einer anderen Aktivität als Input benötigt wird (Flow). Hiervon können die Fälle abgegrenzt werden, in denen mehrere Aktivitäten auf dieselbe Ressource zugreifen (Sharing) oder aber gemeinsam eine Ressource hervorbringen (Fit).

Ressourcen: Ressourcen stellen die Inputs eines Geschäftsprozesses resp. seiner einzelnen Aktivitäten dar und können in den Kategorien Arbeit und Kapital zusammengefasst werden. Beispiele sind menschliche Arbeitskräfte, automatisierte Fertigungsmaschinen, Lagerräume und unfertige Produkte. Zur Prozessmodellierung werden Ressourcen verschiedenen *Aktivitäten zugeordnet* und mit *Attributen versehen* [Tuma1996]. Bspw. kann ein Kanban-Koordinationsmechanismus durch Definition der maximal erlaubten Anzahl unfertiger Produkte in der Ressource „Pufferlager“ abgebildet werden [Schr2004].

4.5 Modellierung der ökonomischen Auswirkungen von IT

Die Klassifizierung ökonomischer Effekte von IT und die Modellierung von deren Auswirkungen auf Prozesskomponenten bestimmt maßgeblich die Wahl der geeigneten Kennzahlen. Während zur Klassifizierung der Performance-Effekte von IT im Allgemeinen und RFID im Speziellen einige Ansätze vorhanden sind, wurden die Auswirkungen von IT-Effekten auf Prozessmodelle bislang nur sporadisch untersucht. Ansätze zur Übertragung von RFID-Effekten auf Prozessmodelle bestehen bislang nicht.

Klassifizierung ökonomischer Effekte von IT: Eine Möglichkeit der Klassifizierung von Performance-Auswirkungen von IT auf Prozesse besteht in der Einteilung in neun Effekte, im Rahmen derer bspw. zwischen der Substitution menschlicher Arbeit („*automational*“), Änderung in der Reihenfolge von Prozessschritten („*sequential*“) und verbesserter Analyse von Informationen („*analytical*“) unterschieden werden kann [Dave1993]. Alternativ hierzu werden IT-Effekte auch oftmals auf die drei Effekte „*Automatisierung*“, „*Informationsverbesserung*“ und „*Prozesstransformation*“ verdichtet [MoGK1996].

Modellierung der Auswirkungen von IT-Effekten auf Prozesse: Bei der Übertragung dieser IT-Effekte auf Prozessmodelle können einzelne Effekte wie „*Automatisierung*“ und „*Sequenzänderung*“ (Transformation) identifiziert werden, deren Auswirkungen (z.B. Reduktion des Ressourcenaufwands) in Prozessmodellen abgebildet werden können. Für andere, besonders beim RFID-Einsatz relevante IT-Effekte wie „*Tracking*“, „*Informationsverbesserung*“ und „*Koordination über Distanz*“ konnte keine geeignete Abbildungsform gefunden werden [Giag1999].

Klassifizierung ökonomischer Effekte von RFID: In Anlehnung an [MoGK1996] unternehmen erste Ansätze der Klassifizierung RFID-bedingter Performance-Effekte eine Dreiteilung in *Automatisierungs-, Informations- und Transformationseffekte* [Tell2006]. Dabei führt der Automatisierungseffekt lediglich zu einer Substitution vormals manuell durchgeführter Informationserfassungstätigkeiten. Etwaige qualitative Verbesserungen der erfassten Informationen werden ausgeschlossen. Bspw. führt die RFID-basierte Palettenerfassung im Wareneingang von Einzelhandelsunternehmen zu keinen nennenswerten Informationsverbesserungen im Vergleich zur bereits sehr robusten Barcode-Technologie. Als Resultat des Informationseffekts kann hingegen die Datenqualität hinsichtlich mindestens einer Dimension verbessert werden (z.B. Korrektheit, Vollständigkeit). Im Ergebnis kann bspw. eine Kostensenkung bei der Behebung von Fehlbeständen im Warenwirtschaftssystem oder ein Rückgang von Out-Of-Stock-Situationen erreicht werden. Im Falle des Transformationseffekts kann eine Leistungssteigerung nicht alleine durch die Technologieeinführung realisiert werden. Als Komplementärinvestition muss der Ausgangsprozess verändert werden. Bspw. kann durch Einführung einer zusätzlichen Datenerfassung zwischen dem Front- und Backstore-Bereich im Einzelhandel und entsprechend angepassten Koordinationsmechanismen die Nachbefüllung von Regalen verbessert werden [Tell2006].

5 Methodische Vorgehensweise

Unter Anwendung der vorgestellten Ansätze aus Performance-Messung und Prozessmodellierung soll das Ziel des Dissertationsvorhabens, die Entwicklung eines prozessbasierten und multidimensionalen Messsystems zur systematischen und anwendungsübergreifenden Erfassung des RFID-Einflusses auf die Leistungsfähigkeit von Logistikprozessen, in fünf Teilschritten realisiert werden.

Phase I: Modellierung von Logistikprozessen und RFID-Effekten

Im ersten Schritt werden zwei verschiedene Zielsetzungen verfolgt. Zunächst werden für die betrachteten Logistikprozesse „Kommissionierung“, „Kanban-Produktion“ und „Warenein-/ausgang“ repräsentative Ausprägungsformen identifiziert und in generische Prozessmodelle überführt. Mit Hilfe des Verfahrens der Prozessdekomposition erfolgt eine schrittweise Disaggregation der Gesamtprozesse ausgehend von abstrakten Darstellungsformen (s. Abschnitt 4.4) hin zu Teilprozessen, bis schlussendlich auf der Mikroebene anwendungsspezifische Aktivitäten, deren Anordnung sowie Abhängigkeiten zwischen Ressourcen und Aktivitäten identifiziert werden können. Prozessdekomposition ist ein vielfach in der Praxis angewendetes Verfahren, das bspw. als Teil des „Toyota Production System“ wesentlich zum Verständnis von Prozessen in der Automobilindustrie und deren kontinuierlicher Verbesserung beiträgt [SpBo1999]. Zudem werden mögliche Auswirkungen der drei identifizierten RFID-Performance-Effekte auf Prozesse im Allgemeinen kategorisiert. Dieser Einfluss kann sich bspw. in der Reduktion des Ressourcenaufwandes einzelner Aktivitäten, der Veränderung von Abhängigkeiten und dem Wegfallen bzw. Hinzukommen von Aktivitäten manifestieren. Im Falle des Automatisierungseffekts betreffen diese Veränderungen definitionsgemäß lediglich eine *Reduktion der aufgewendeten Ressourcen* zur Informationserfassung. Im Wesentlichen wird es sich hierbei um Personalkostenreduktionen handeln. Zur Bestimmung der Auswirkungen qualitativ verbesserter Informationen (Informationseffekt) wird in Anlehnung an [Tell2006] der Ansatz der *Qualitätskosten* [HeRe2004] herangezogen. Es soll gezeigt werden, dass neben einer Veränderung des Ressourceneinsatzes (Kosten zur Informationserfassung, Fehlerbehebung etc.) auch *Aktivitäten* entfallen bzw. hinzukommen können. So kann eine qualitativ verbesserte Informationslage bspw. die Einführung zusätzlicher Datenerfassungsaktivitäten (*Präventionskosten*) notwendig und gleichzeitig routinemäßige Fehlerkontrollen (*Kontrollkosten*) überflüssig machen. Im Falle des Transformationseffekts verändern sich sowohl Ressourceneinsatz und Aktivitäten als auch

dahinterstehende Abhängigkeiten. Mögliche Änderungen dieser Elemente sollen mit Hilfe der *Theorie der Komplementaritäten* [MiRo1992], anhand derer das optimale Ausmaß einer Prozessänderung als Folge *komplementärer* und *konkurrierender* Zielsetzungen bestimmt werden kann [BrRA1997], abgeleitet werden.

Als Ergebnis der ersten Phase können einerseits „typische“ Aktivitäten, Ressourcen und Abhängigkeiten repräsentativer Kommissionierungs-, Kanban- und Warenein- und -ausgangsprozesse bestimmt werden. Andererseits wird gezeigt, welche Veränderungen RFID-Technologie an abstrakten Prozess(-element)en bewirken kann.

Phase II: ABC-Ansatz zur Erfassung des Automatisierungseffekts

Daran anknüpfend sollen diejenigen Aktivitäten (bzw. deren Ressourceneinsatz) in den betrachteten Prozessmodellen identifiziert werden, welche aufgrund des Automatisierungseffekts prinzipiell beeinflusst werden können. Unter der Berücksichtigung von ökonomischen (z.B. Betriebskosten des RFID-Systems) und technologischen Entscheidungsgrößen (z.B. Auslesegenauigkeit des RFID-Systems) werden Kennzahlen abgeleitet, die auf Aktivitätenebene die Veränderung des Ressourceneinsatzes quantifizieren können. Durch Integration dieser Kennzahlen in ein ABC-Modell wird die Grundlage für das zu entwickelnde Messsystem gelegt. Zudem werden verschiedene Möglichkeiten der Erhebung dieser Kennzahlen aufgezeigt (ERP-Datenauswertung, Mitarbeiterinterviews, Zeitmessungen etc.).

Phase III: Ausdehnung des ABC-Ansatzes auf die Auswirkungen des Informationseffekts

Im dritten Schritt erfolgt die Identifikation derjenigen Aktivitäten und Ressourcen aus den Prozessmodellen, die durch eine verbesserte Datenqualität (Informationseffekt) prinzipiell beeinflusst werden können. Analog zum Vorgehen in Phase II werden unter Berücksichtigung unterschiedlicher Determinanten (z.B. Datengranularität⁷) Kennzahlen identifiziert, die einen veränderten Ressourceneinsatz und zusätzlich das Hinzukommen resp. Entfallen von Aktivitäten quantitativ erfassen können. Das Ergebnis der Phasen II und III ist ein Modell zur aktivitätsbasierten Performance-Messung aufgrund von Automatisierung und verbesserter Informationslage beim RFID-Einsatz. Dieses Modell kann allerdings nur Teilaspekte des Informationseffekts messen und muss in Phase IV erweitert werden.

⁷ Es ist davon auszugehen, dass bspw. die erhöhte Datengranularität bei Erfassung auf Item-Level weitreichendere Veränderungen von Ressourceneinsatz und Aktivitäten bedingt als die Erfassung auf Behälterebene.

Phase IV: Integration strategieorientierter Kennzahlen in das Messsystem

Während zur Messung des Automatisierungseffekts die Erfassung aktivitätsbasierter (finanzieller) Kennzahlen ausreichend ist, gewinnen mit der Nutzung des *Informationseffekts* weitere, über eine alleinige Kostensenkung hinausgehende *strategische Zielsetzungen* zunehmend an Bedeutung. So lassen sich qualitative Zielgrößen wie Liefertreue, Flexibilität, Warenverfügbarkeit oder Fehlerhäufigkeit meist nicht unmittelbar auf der Ebene einzelner Aktivitäten in Form von Kosten quantifizieren (obgleich selbstverständlich auch diese Zielgrößen Umsatz und (Opportunitäts-)Kosten eines Unternehmens beeinflussen). Diese Größen müssen auf Ebene des Gesamtprozesses in Form nicht-finanzieller Kennzahlen erhoben und in das Messmodell integriert werden, um den Informationseffekt vollständig zu erfassen. Eine große Auswahl nicht-finanzieller Kennzahlen bietet bspw. [Phil2000]. Die Ableitung der Kennzahlen aus Strategiezielen und Integration in das Messsystem erfolgt mit Hilfe des *Auditing and Enhancing Framework* [MeSt2000].

Zur Erfassung der Auswirkungen des *Transformationseffekts* muss beachtet werden, dass eine Messung auf Ebene der Prozesselemente bei weitreichenden Veränderungen von Aktivitäten, Ressourcen und Abhängigkeiten keine sinnvollen Vergleichswerte mehr liefert. Dies macht die Integration *prozessbasierter* finanzieller (z.B. Gesamtkosten des Prozesses) und nicht-finanzieller (z.B. Durchlaufzeit) Kennzahlen erforderlich. Die Identifikation relevanter Kennzahlen erfolgt durch Ableitung aus verschiedenen *strategischen Gesamtzielen* der Prozessumgestaltung [Phil2000] [MeSt2000] [Dave1993]. Das Ergebnis der Phasen I – IV ist ein Messsystem, das in Abhängigkeit von der Art der angestrebten Leistungssteigerung mit RFID-Technologie ein Set von Kennzahlen zur Performance-Messung zur Verfügung stellt.

Phase V: Anwendung des Messsystems in realen Logistikanwendungen

In der letzten Phase erfolgt eine Anwendung des Messsystems in realen Logistikszenarien. Die Erfassung der für die jeweiligen Prozesse identifizierten Kennzahlen erfolgt zu verschiedenen Zeitpunkten vor und nach der RFID-Einführung. Einerseits sollen die Auswirkungen der RFID-Technologie auf die Prozessleistung sichtbar gemacht werden. Andererseits sollen durch die Anwendung multivariater Analyseverfahren⁸ Zusammenhänge zwischen den für die RFID-Technologie relevanten Leistungsdimensionen sichtbar gemacht und Aussagen hinsichtlich der

⁸ Angestrebt ist bspw. eine Dimensionsreduktion auf einzelne relevante „Meta-Kennzahlen“ durch Anwendung der Faktorenanalyse, anhand derer der Einfluss der RFID-Technologie sichtbar gemacht werden kann.

„optimalen“ Einbindung von RFID-Technologie in die Prozessgestaltung und Organisationsstruktur von Unternehmen getroffen werden.

6 Zusammenfassung

Die Messung des ökonomischen Nutzens der RFID-Technologie steht noch am Anfang. Alternativ zu bisherigen Methoden der Performance-Messung beim IT-Einsatz wird in dem Dissertationsvorhaben ein prozessbasierter Ansatz verfolgt, der über die Erfassung von Kosteneinsparungen aufgrund von Automatisierung hinausgeht. Auf Basis der grundlegenden Unterscheidung zwischen Prozess- und Aktivitätenebene soll ein Kennzahlensystem entwickelt werden, das sowohl kostenorientierte als auch finanziell schwer greifbare strategische Zielgrößen, die aufgrund verbesserter Informationsqualität und möglicher Prozesstransformationen realisiert werden können, in einem angemessenen Verhältnis repräsentiert. Die Anwendung des Messsystems in realen Logistikanwendungen soll Erkenntnisse über das quantitative Ausmaß der verschiedenen Performance-Effekte von RFID liefern und grundlegende Leistungsdimensionen der Technologie aufdecken.

Literaturverzeichnis

- [Aber2006] Aberdeen Group: ROI in RFID. Report, Boston 2006.
- [Ange2005] Angeles, R.: RFID Technologies – Supply-Chain Applications and Implementation Issues. In: Information Systems Management 22/1 (2005), S. 51-65.
- [BoDL2004] Bone, T.; Dirkling, S.; Lammers, W.: Erwartungen bei Handel und Industrie. In: ten Hompel, M.; Lange, V. (Hrsg.): Radio Frequenz Identifikation 2004 – Logistiktrends für Industrie und Handel, Dortmund 2004, S. 34-83.
- [BoFW2003] Bourne, M.; Franco, M.; Wilkes, J.: Corporate Performance Measurement. In: Measuring Business Excellence 7/3 (2003), S. 15-21.
- [BrHi1998] Brynjolfsson, E.; Hitt, L.M.: Beyond the Productivity Paradox. In: Communications of the ACM 41/8 (1998), S. 49-55.
- [BrRA1997] Brynjolfsson, E.; Renshaw, A.; van Alstyne, M.: The Matrix of Change: A Tool for Business Process Reengineering. In: Sloan Management Review 38/2 (1997), S. 37-54.

- [Carr2004] Carr, N.: Does IT Matter? Boston 2004.
- [Chan1977] Chandler, A.: The Visible Hand – Managerial Revolution in American Business. Boston 1977.
- [ChSh2000] Chand, G.; Shirvani, B.: Implementation of TPM in Cellular Manufacture. In: Journal of Materials Processing Technology 103/1 (2000), S. 149-154.
- [Dave1993] Davenport, T.: Process Innovation – Reengineering Work through Information Technology. Boston 1993.
- [DeGK2003] Dedrick, J.; Gurbaxani, V.; Kraemer, K.: Information Technology and Economic Performance – A Critical Review of the Empirical Evidence. In: ACM Computing Surveys 35/1 (2003), S. 1-28.
- [Ditt2006] Dittmann, L.: Der angemessene Grad an Visibilität in Logistik-Netzwerken. Dissertation, St. Gallen 2006.
- [FICD2005] Fleisch, E.; Christ, O.; Dierkes, M.: Die betriebswirtschaftliche Vision des Internets der Dinge. In: Fleisch, E.; Mattern, F. (Hrsg.): Das Internet der Dinge – Ubiquitous Computing und RFID in der Praxis. Berlin 2005, S. 3–37.
- [GhNo1996] Ghalayini, A.; Noble, J.: The Changing Basis of Performance Measurement. In: International Journal of Operations and Production Management 16/8 (1996), S. 63-80.
- [Giag1999] Giaglis, G.: Modelling the Impact of Information Systems in Business Process Simulation Models. In: Ades, M. (Hrsg.): Proceedings of the 1999 Advanced Simulation Technologies Conference, San Diego 1999, S. 172-177.
- [GiHa2006] Gillert, F.; Hansen, W.-R.: RFID für die Optimierung von Geschäftsprozessen. München 2006.
- [Gros1993] Grossman, E.: How to Measure Company Productivity. In: Christopher, W.; Thor, C. (Hrsg.): Handbook for Productivity Measurement and Improvement, Cambridge (USA) 1993.
- [HeRe2004] Heizer, J.; Render, B.: Operations Management. 7. Auflage, Upper Saddle River 2004.
- [HoDe2005] ten Hompel, M.; Deuse, J.: Ganzheitliche Bewertung und Optimierung manueller Arbeitsplätze in Kommissioniersystemen. Schlussbericht des AiF-Forschungsvorhabens Nr. 13904, Dortmund 2005.
- [Ibol2005] Ibold, C.: Prozessoptimierung durch RFID. In: Logistik inside online 2005, <http://www.logistik-inside.de/fm/2248/Ibold.pdf> (Abruf am 25.07.2006).
- [Info2006] Informationsforum RFID (Hrsg.): RFID – Leitfaden für den Mittelstand. Berlin 2006.

- [KaNo1992] Kaplan, R.; Norton, D.: The Balanced Scorecard – Measures that Drive Performance. In: Harvard Business Review 70/1 (1992), S. 71-79.
- [KaCo1998] Kaplan, R.; Cooper, R.: Cost & Effect – Using Integrated Cost Systems to Drive Profitability and Performance. Boston 1998.
- [KaAn2003] Kaplan, R.; Anderson, S.: Time-Driven Activity-Based Costing. Harvard Business School Paper, Boston 2003.
- [KeEJ1989] Keegan, D.; Eiler, R.; Jones, C.: Are your Performance Measures Obsolete? In: Management Accounting 70/12 (1989), S. 45-50.
- [Laub2006] Laubacher, R.; Kothari, S.P.; Malone, T.W.; Subirana, B.: What is RFID Worth to Your Company? – Measuring Performance at the Activity Level. MIT Sloan School Paper # 224, Cambridge (USA) 2006.
- [Malo1999] Malone, T.W.; Crowston, K.; Lee, J.; Pentland, B.; Dellarocas, C.; Wyner, G.; Quimby, J.; Osborn, C.S.; Bernstein, A.; Herman, G.; Klein, M.; O'Donnell, E.: Tools for Inventing Organizations: Toward a Handbook of Organizational Processes. In: Management Science 45/3 (1999), S. 425–443.
- [MeSt2000] Medori, D.; Steeple, D.: A Framework for Auditing and Enhancing Performance Measurement Systems. In: International Journal of Operations and Production Management 20/5 (2000), S. 520-533.
- [MiRo1992] Milgrom, P.; Roberts, J.: Economics, Organization and Management. London 1992.
- [MoGK1996] Mooney, J.; Gurbaxani, V.; Kraemer, K.: A Process Oriented Framework for Assessing the Business Value of Information Technology. In: The DATA Base for Advances in Information Systems 27/2 (1996), S. 68-81.
- [Muel2003] Müller, G.; Kreutzer, M.; Strasser, M.; Eymann, T.; Hohl, A.; Nopper, N.; Sackmann, S.; Coroama, V.: Geduldige Technologie für ungeduldige Patienten: Führt Ubiquitous Computing zu mehr Selbstbestimmung? In: Mattern, F. (Hrsg.): Total Vernetzt, Heidelberg 2003.
- [Muel2006] Müller, G.: Privacy and Security in Highly Dynamic Systems. In: Communications of the ACM 49/9 (2006), S. 28-31.
- [Naka1988] Nakajima, S.: Introduction to TPM. Cambridge (USA) 1988.
- [Neel1999] Neely, A.: The Performance Measurement Revolution: Why now and what next? In: International Journal of Operations and Production Management 19/2 (1999), S. 205-228.
- [Park2000] Parker, C.: Performance Measurement. In: Work Study 49/2 (2000), S. 63-66.

- [Pete2000] Petersson, P.: Process Efficiency and Capacity Flexibility – Developing a Support Tool for Capacity Decisions in Manual Assembly Systems. Dissertation, Linköping 2000.
- [Phil2000] Phillips, J.: The Consultant's Scorecard – Tracking Results and Bottom-Line Impact of Consulting Projects. New York 2000.
- [RoVo2006] Rob, P.; Voerste, A. (Hrsg.): METRO Group RFID Newsletter 04/2006, Düsseldorf 2006.
- [Schr2004] Schroer, B.: Simulation as a Tool in Understanding the Concepts of Lean Manufacturing. In: Simulation 80/3 (2004), S. 171-175.
- [Schu1999] Schulte, C.: Logistik. 3. Auflage, München 1999.
- [SiZa1995] Sinclair, D.; Zairi, M.: Effective Process Management through Performance Management. In: Business Process Re-engineering and Management Journal 1/1 (1995), S. 75-88.
- [SpBo1999] Spear, S.; Bowen, K.: Decoding the DNA of the Toyota Production System. In: Harvard Business Review 77/5 (1999), S. 97-106.
- [StGi2006] Strüker, J.; Gille, D.: RFID in Deutschland – Status quo, Chancen und Hürden aus Unternehmenssicht. Institut für Informatik und Gesellschaft, Universität Freiburg 2006.
- [Stra2005] Strassner, M.: RFID im Supply Chain Management – Auswirkungen und Handlungsempfehlungen am Beispiel der Automobilindustrie. Wiesbaden 2005.
- [Subi2003] Subirana, B.; Eckes, C.; Herman, G.; Sarma, S.; Barrett, M.: Measuring the Impact of Information Technology on Value and Productivity using a Process-Based Approach: The case for RFID Technologies. CCS Paper # 223, MIT Sloan School, Cambridge (USA) 2003.
- [Tang2002] Tangen, S.: Understanding the Concept of Productivity. In: Proceedings of the 7th Asia Pacific Industrial Engineering and Management Systems Conference, Taipei 2002.
- [Tang2004] Tangen, S.: Evaluation and Revision of Performance Measurement Systems. Dissertation, Stockholm 2004.
- [Tell2006] Tellkamp, C.: The Impact of Auto-ID Technology on Process Performance – RFID in the FMCG Supply Chain. Dissertation, St. Gallen 2006.
- [Tuma1996] Tumay, K.: Business Process Simulation. In: Charnes, J.; Morrice, D.; Brunner, D.; Swain, J. (Hrsg.): Proceedings of the 1996 Winter Simulation Conference, Coronado (USA) 1996, S. 93-98.

Softwareagenten in der Krankenhauslogistik

Ein Ansatz zur effizienten Ressourcenallokation

Dipl.-Wirtsch.Inf Christoph Niemann
Lehrstuhl für Wirtschaftsinformatik (BWL VII)
Universitätsstr. 30
95447 Bayreuth
christoph.niemann@uni-bayreuth.de

betreut von Prof. Dr. Torsten Eymann
Lehrstuhl für Wirtschaftsinformatik (BWL VII)
Universitätsstr. 30
95447 Bayreuth
torsten.eymann@uni-bayreuth.de

Abstract

Die Dissertation führt Softwareagenten in die Krankenhauslogistik zur Unterstützung der Terminplanung ein. Die Softwareagenten agieren in einem logischen Abbild der Realität, das durch die Anwendung von Ubiquitous-Computing-Technologien mit der Wirklichkeit synchronisiert wird. Auf dieser informationstechnischen Grundlage verhandeln die Softwareagenten auf einem künstlichen Markt, um so den Terminplan des Krankenhauses zu optimieren. Der Fokus der Arbeit liegt dabei insbesondere auf der Anwendung des Paradigmas der Softwareagenten zur Umsetzung der Vision des Ubiquitous Computing und den daraus resultierenden Effekten. Die Evaluation der Arbeit erfolgt anhand der realisierten Effizienz und Fairness.

1 Motivation

In der Vision des „Ubiquitous Computing“ (UC) [Wei91] postuliert Weiser als Folge der fortwährenden Miniaturisierung und der Durchdringung der Wirklichkeit mit Informationstechnik die mögliche Automatisierung von Entscheidungen im Hintergrund ohne Zutun des Menschen. UC verspricht damit eine „calm technology“ [Wei98], die durch die Verknüpfung von allge-

genwärtiger Sensorik (und Effektorik) mit praktisch unbegrenzter Verarbeitungskapazität in die Lage versetzt wird, die Wirklichkeit detailgetreu abzubilden. Die dabei anfallenden Daten können verarbeitet werden und bilden die Grundlage für vollständig automatisierte Entscheidungssysteme, in die der Mensch nicht mehr eingreifen muss.

Ausgehend von dieser Vision wendet diese Arbeit UC-Technologien in einer Domäne an, die diese vollständige Automatisierung von Entscheidungen aus prinzipiellen Gesichtspunkten *nicht* ermöglicht: Im Krankenhausumfeld können Informationssysteme (IS) zur Unterstützung der Planung eingesetzt werden, die Entscheidungen können jedoch nicht ohne medizinisches Fachwissen getroffen werden, welches bisher nur von Menschen eingebracht werden kann. Um solches Wissen im Entscheidungsprozess berücksichtigen zu können, ist daher die Kombination von maschineller Information und menschlichem Wissen notwendig. Die Entscheidungskompetenz liegt weder ausschließlich bei den IS, noch allein bei den menschlichen Akteuren. Nur das Zusammenspiel kann zu sinnvollen Ergebnissen führen. Die Intelligenz eines Gesamtsystems ist damit eine Form der „hybriden Intelligenz“ [Ram03]. Die dieser Arbeit zu Grunde liegende Fragestellung ist damit die nach der Natur dieser hybriden Intelligenz. Insbesondere untersucht die Arbeit Ansätze, mit denen hybride Intelligenz realisiert werden kann, ohne auf die von Weiser postulierte ruhige Technologie verzichten zu müssen.

2 Zielsetzung der Arbeit

Zielsetzung der Dissertation ist die Konzeption und Entwicklung eines Systems zur flexiblen Ressourcenplanung am Beispiel der Domäne Krankenhaus. Dabei wird untersucht, ob sich ein Multiagentensystem (MAS), in dem Softwareagenten auf Märkten über die Ressourcenzuweisung verhandeln [Eym01; Wel96], eignet, eine effiziente Ressourcenallokation des Gesamtsystems zu erreichen. Die verteilte Entscheidungskompetenz zwischen Mensch und Maschine stellt die wichtigste Fragestellung der Arbeit dar. Aus ihr leitet sich letztendlich ab, ob sich UC-Technologien auch in Problem-domänen eignen, deren Wissen nicht vollständig automatisiert zu erfassen ist.

Die Frage, wie Information, die nicht durch die immer stärker miniaturisierte IT autonom erfasst werden kann, in ein IS eingebracht werden kann, ohne die menschlichen Akteure durch Abfrage solcher Information zu überlasten, bildet den Kern der Arbeit. Ihre Untersuchung wird in der Domäne der Kliniken durchgeführt, ist jedoch nicht grundsätzlich von ihr abhängig. Sie

lässt sich allgemein auf Problemfelder abstrahieren, die mit verteilter Entscheidungskompetenz umgehen müssen, da nicht alle Daten maschinell erfassbar sind.

Die Arbeit leistet auf zwei Ebenen Beiträge zur Forschung:

- *Artefakt*: Auf der Ebene der konkreten Instanz soll das Artefakt die Ressourcenallokation zufrieden stellend lösen. Im Artefakt wird der vorgeschlagene Ansatz zur Problemlösung implementiert und ermöglicht so die Simulation des Problems im Modellmaßstab. Erst wenn sich das Artefakt in diesem Maßstab bewährt, kann es in größeren Maßstäben evaluiert werden. Zur Bewertung des Nutzens dient die Effizienz der Planung.
- *Methoden*: Methoden liegen auf einer höheren Abstraktionsebene als Artefakte und sind nicht an ein spezifisches Artefakt gebunden. Sie lassen sich auf verschiedene Artefakte anwenden, die dem gleichen Problemfeld zugeordnet sind. Damit abstrahieren sie von Eigenschaften eines spezifischen Artefaktes und erlauben den Vergleich verschiedener Artefakte untereinander. In dieser Arbeit wird das Artefakt an Hand der erreichten Effizienz bewertet. Die Nutzenbewertung setzt eine Metrik voraus, in der die verschiedenen quantitativen Größen zur Zielgröße der Effizienz aggregiert werden können. Die Entwicklung dieser Metrik ist ein Ziel der Arbeit.

Neben der Effizienz-Bewertung des Artefaktes steht in der Arbeit die Untersuchung der Effekte des UC im Vordergrund. Durch den Einsatz verschiedener UC-Technologien ergibt sich die Möglichkeit, die Auswirkungen von UC in einer begrenzten Domäne zu untersuchen. Insbesondere die Frage nach dem Umschlagspunkt des Gesamtsystems zwischen von Ordnung und Chaos ist hier interessant. Das eingesetzte MAS ermöglicht die Verarbeitung der anfallenden Daten, kann aber bei zu starker Automatisierung dazu führen, dass das Gesamtsystem nicht mehr in der Lage ist, die notwendige Flexibilität aufrecht zu erhalten. Das System wird zu starr in der Entscheidungsfindung. Auf der anderen Seite muss das System die Datenmenge genügend stark reduzieren, um die Menschen nicht mit „information overload“ [Mae94] zu überlasten. Diese Arbeit untersucht diese Grenze zwischen Ordnung und Chaos.

Schließlich stellt die Übertragbarkeit der Ergebnisse eine wichtige Zielgröße dar. Von der Domäne des Krankenhauses soll auf andere Domänen verallgemeinert werden, deren Entscheidungsfindung sich – wie in Krankenhäusern – nicht vollständig automatisieren lässt. Aus dem von Mertens postulierten Ziel der Wirtschaftsinformatik, der „sinnhaften Vollautomation“

[Mer95], wird somit insbesondere untersucht, wie sich der Bestandteil der „Sinnhaftigkeit“ in der Domäne darstellen kann.

2.1 Related Work

Die Arbeit entwickelt das EMIKA-Modell (*Echzeitgestützte Mobile Terminplanung in Krankenhausanwendungen*) weiter, das die Terminplanung in Krankenhäusern unterstützt [EMS06; NE06; SE04; MKS+03]. EMIKA ermöglicht die Substitution einer einzelnen Ressource innerhalb eines Behandlungstermins, beschränkt sie jedoch auf die Klasse der Ärzte. Andere Ressourcen, die Leistungen im Krankenhaus anbieten, werden nicht berücksichtigt. Das Projekt führt den Marktmechanismus als Koordinationsinstrument ein, beschränkt die Verhandlungen jedoch auf ein begrenztes Feld. Diese Dissertation führt die Forschung hier weiter und bettet EMIKA in ein theoretisches Fundament ein, um Problemstellungen des Ubiquitous Computing untersuchen zu können.

Im Rahmen des abgeschlossenen Schwerpunktprogramms 1083 der DFG existieren weitere Projekte deren Forschungsschwerpunkte denen dieser Arbeit ähneln. Dazu gehören insbesondere die Projekte *Policy-Agents* [BC06] und *MedPage* [PZR+06]. Beide Projekte befassen sich mit der Terminplanung bzw. Koordination klinischer Prozesse, wobei *Policy-Agents* sich auf die Terminplanung für Operationsräume beschränkt, während *MedPage* den gesamten Behandlungsprozess umfasst. *MedPage* legt als Nutzenmaß für die abschließende Evaluation so genannte „years of well being“ zu Grunde und stellt damit auf eine medizinische Maßzahl zur Bewertung ab. EMIKA rückt ökonomische Maße in den Vordergrund, die medizinische Qualität der Behandlungen wird nicht beeinflusst. Sowohl *Policy-Agents* als auch *MedPage* gehen im Gegensatz zu EMIKA von der vollständigen Automatisierung der Entscheidungen aus und blenden damit die Interaktion zwischen Mensch und Maschine bewusst aus. Diese Interaktion stellt einen wichtigen Schwerpunkt in EMIKA dar. Keines der beiden Projekte implementiert mobile Endgeräte als Kommunikationskanal zwischen Agent und dem von ihm abgebildeten Prinzipal.

Das Projekt *ChariTime* [MG01] stellt einen Ansatz zur Terminplanung in Krankenhäusern auf Basis eines MAS vor, abstrahiert jedoch auf die Funktionsebene. Damit werden nicht mehr einzelne Personen als Akteure betrachtet, sondern einzelne Abteilungen im Krankenhaus.

3 Forschungsmethode

Diese Arbeit nutzt zur Untersuchung der Effekte des UC „design science“ [HMPR04]. Im Vordergrund steht die Lösung des Terminplanungsproblems in Krankenhäusern. Über den Weg der Innovation soll ein Artefakt geschaffen werden, welches das Problem löst und das sich bezüglich seines Nutzens evaluieren lässt.

Über die beiden Schritte „build“ und „evaluate“ [HMPR04] gliedert sich die Arbeit in zwei große Teilabschnitte auf: Im ersten Schritt entsteht das Artefakt und die theoretische Fundierung der beiden angewandten Technologien wird erläutert. Das Artefakt liegt auf der untersten Ebene der von Vaishnavi und Kuechler vorgeschlagenen Kategorisierung der Ergebnisse [VK05] von design science und bildet die Grundlage für weitergehende Untersuchungen.

Im zweiten Schritt wird das Artefakt an Hand seines Nutzens bewertet. Da der Nutzen nicht offensichtlich quantifizierbar ist, entwickelt die Arbeit zu diesem Zweck eine geeignete Metrik. Diese Metrik stellt ökonomische Maßzahlen in den Vordergrund, da durch den Einsatz des Artefaktes Veränderungen der Effizienz und nicht der medizinischen Behandlungsqualität zu erwarten sind. Unabhängig von EMIKA werden die gleichen Behandlungen durchgeführt, die auch ohne das System ausgeführt würden. Deshalb bleibt die medizinische Behandlungsqualität konstant und ändert sich allenfalls mittelbar in Form von veränderter Patientenzufriedenheit durch kürzere Behandlungsdauern. Diese Effekte lassen sich jedoch im Modellmaßstab nicht untersuchen, so dass sie ausgeblendet werden.

Die Metrik stellt ein Ergebnis auf der Ebene der Methoden [VK05] dar, die notwendig sind, um Artefakte zu bewerten. Auf dem Artefakt und der Metrik aufbauend, können die zu Grunde liegenden Prinzipien des Marktmechanismus und die daraus entstehenden Phänomene des UC untersucht werden, was den dritten Teil der Arbeit darstellt.

4 Theoretische Grundlagen

Dieser Abschnitt stellt im ersten Teil die Grundlagen der Problemdomäne dar und erläutert ihre besonderen Eigenschaften, die die Anwendung von UC zu einem interessanten Forschungsgebiet werden lassen. Der zweite Abschnitt geht auf UC ein und erläutert die eingesetzten Technologien.

4.1 Domäne Krankenhaus

Kliniken stehen unter erheblichem Kostendruck. Mit Einführung von Fallpauschalen zur Abrechnung von Leistungen hat sich die Einnahmenstruktur von Krankenhäusern grundlegend verändert, da nicht mehr Liegetage eines Patienten, sondern Diagnosen die Grundlage für eine pauschalierte Bezahlung bilden. Um die dadurch verursachte Reduzierung der Einnahmen zu kompensieren, müssen Kliniken versuchen, ihre Kosten mindestens in gleichem Maße zu verringern, um wettbewerbsfähig (und damit überlebensfähig) zu bleiben.

Um die bestmögliche Behandlung der Patienten sicherzustellen, durchlaufen diese standardisierte Behandlungspfade im Sinne der „evidenzbasierten Medizin“ [FYLA04]. Solche Behandlungspfade umfassen die nach den „best practices“ zusammengestellten Behandlungsschritte bezogen auf die gestellte Diagnose. Die Pfade bilden damit den gesamten Krankenhausaufenthalt eines Patienten ab und verhalten sich analog zu – wenn auch stark reglementierten – Geschäftsprozessen in Unternehmen. Ein Pfad umfasst den sequentiellen Ablauf der aufeinander folgenden Untersuchungen. Da Behandlungen nicht immer streng sequentiell ablaufen, sehen Behandlungspfade die Möglichkeit von Verschiebungen einzelnen Behandlungsschritten oder Fallauswahlen vor. Weiterhin können sie Ausnahmebehandlungen abbilden und erlauben die flexible Anpassung bei sich verändernden Diagnosen. Die Planung in Krankenhäusern gewinnt durch Behandlungspfade an Struktur, wodurch eine stationsübergreifende Planung erst ermöglicht wird.

Die Allokationsplanung von Ressourcen geschieht gegenwärtig vornehmlich manuell durch menschliche Bearbeiter. Diese Form der Allokation skaliert jedoch nur bedingt mit zunehmender Größe des Allokationsproblems und der damit einhergehenden Vergrößerung des potenziellen Lösungsraumes. Die menschliche Allokation erreicht daher zunehmend ihre Grenzen [MKS+03, S. 12]. Das in Kliniken vorhandene Planungspotenzial soll diese Arbeit durch die Umsetzung der Entscheidungsunterstützung mittels UC-Technologien lösen.

4.2 Ubiquitous Computing

UC bildet ein Paradigma der Informationsverarbeitung, das die Miniaturisierung der IT als Voraussetzung für die weitgehende Durchdringung der Wirklichkeit mit Informationssystemen darstellt. UC lässt sich in zwei grundlegende Bestandteile, die *Allgegenwärtigkeit* (engl. Ubiquity) und die *Verarbeitung* (engl. Computing) aufteilen. Diese Dissertation nutzt bestehenden Technologien, um beide UC-Bestandteile zu implementieren. Um die Ergebnisse der Verarbeitung

wieder in die Wirklichkeit zu kommunizieren, werden als dritte Komponente Effektoren benötigt, mit denen die Wirklichkeit entsprechend der Ergebnisse der Verarbeitung beeinflusst werden kann. Den Kreislauf von Sensoren, Verarbeitung und Effektoren zeigt Abbildung 1. Im Diagramm ist auch die Aufteilung in eine physische Welt und deren logische Abbildung zu erkennen. Beide Welten werden durch die Sensorik konsistent gehalten.

4.2.1 Allgegenwärtigkeit durch RFID

Im UC durchdringt miniaturisierte IT die Umgebung und wird so in die Lage versetzt, die Wirklichkeit durch Sensoren wahrzunehmen. Es steht mehr Information zur Verfügung, um Entscheidungen zu unterstützen oder sogar autonom zu treffen. Diese Arbeit setzt *Radio Frequency Identification* (RFID) als berührungslose Funkschnittstelle zur Lokalisierung und Identifizierung von einzelnen Akteuren in Krankenhäusern ein. Die Identität und die Position der Akteure bildet die Grundlage für weitergehende Entscheidungsunterstützung die durch den zweiten Bestandteil des UC verwirklicht wird.

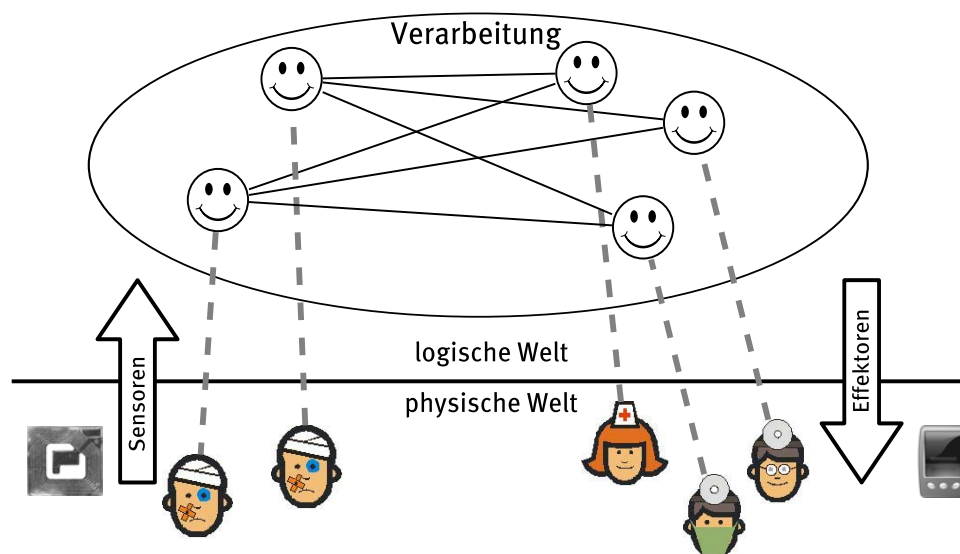


Abbildung 1: Kreislauf von Sensoren, Verarbeitung und Effektoren im EMIKA-System. In der logischen Welt agieren Softwareagenten, in der physischen die menschlichen Akteure. Über die Sensorik wird die Konsistenz beider Welten gesichert.

RFID bietet sich als Technologie in der Klinikdomäne an, da die Kostenstruktur die relativ günstige Einführung ermöglicht: Der Hauptteil der zu veranschlagenden Kosten liegt bei den Lesegeräten, die in verhältnismäßig geringer Zahl benötigt werden. Die Anschaffungskosten der

Transponder, die in großer Anzahl notwendig sind, um alle Akteure auszustatten, liegen im zweistelligen Centbereich. Durch den Einsatz von Polymerdruckverfahren werden sie in naher Zukunft wahrscheinlich noch deutlich sinken [CF03, S. 66].

Die Arbeit abstrahiert von den der RFID-Technologie eigenen Schwächen: Die von RFID verwendete Strahlung wird von Metall reflektiert und von Flüssigkeiten absorbiert [Fin02]. Beide Materialien kommen in Krankenhäusern häufig vor, so dass RFID als Lokalisierungstechnologie zwar im Modellmaßstab geeignet ist, im Krankenhaus jedoch unter Umständen nicht einsetzbar ist. Diese Probleme blendet die Arbeit bewusst aus, da das System nicht an die spezifische RFID-Technologie gebunden ist. Voraussetzung für die zweite UC-Komponente, die Verarbeitung, ist lediglich die verlässliche Lokalisierung der Akteure. Alternative Techniken, wie die WLAN-Lokalisierung haben anderen spezifische Vor- und Nachteile, können RFID jedoch ersetzen, falls sich RFID in Krankenhäusern aus technischen Gründen als nicht geeignet erweist.

4.2.2 Verarbeitung durch Multiagentensysteme

Die Verarbeitungskomponente des UC setzt diese Arbeit mit Multiagentensystemen (MAS) um. MAS bieten große Flexibilität in der Verarbeitung [Kir06] und eignen sich daher zur Verarbeitung in dynamischen Systemen, in denen sich die Anzahl der aktiven Akteure beständig verändert. In Krankenhäusern variiert die Anzahl der Patienten immer wieder: Es werden neue Patienten aufgenommen, während andere entlassen werden.

In MAS verläuft die Informationsverarbeitung (IV) dezentral; jeder Softwareagent handelt auf Grund des ihm eigenen Weltbildes, das durch Sensordaten aktualisiert werden kann. Ein Agent muss dafür Sorge tragen, dass eingehende Sensordaten mit seinem Weltbild abgeglichen werden und er so die zur Entscheidungsunterstützung notwendige Information in seinem Weltbild zur Verfügung hat.

Als Koordinationsmechanismus für die dezentrale IV nutzt diese Arbeit einen Marktmechanismus, in dem unterschiedliche Akteure um knappe Güter (die zu Behandlungen notwendigen Ressourcen) handeln. Ein Geldsubstitut [SE04] dient als Währung auf dem Markt und stellt sicher, dass Patienten, die dringend behandelt werden müssen oder die schon sehr lange auf eine Behandlung warten, in der Lage sind, alle Ressourcen auf dem Markt zu kaufen. Patienten, deren Behandlung aus medizinischer Sicht eine niedrige Priorität hat, müssen unter Umständen Wartezeiten in Kauf nehmen, in denen sie das benötigte Kapital „ansparen“, um dann behandelt zu werden. Auf der Seite der Ressourcenanbieter sichert das Geldsubstitut ab, dass Ärzte nicht

zwei zeitgleiche Behandlungen durchführen müssen. Durch die Preise und Budgets der Agenten fließt medizinisches Fachwissen in den Entscheidungsprozess ein: Über die Höhe des Anfangsbudgets lassen sich verschiedene Behandlungen priorisieren. Diese Transformation von nicht explizitem Wissen in Preise und Budgets ist eine starke Informationsverdichtung, die eine Abwägung zwischen den beiden Polen der ruhigen Technologie und der hybriden Intelligenz darstellt. Die nicht maschinell erfassbare Information wird manuell dem System hinzugefügt, ohne die menschlichen Akteure allerdings mit Informationsnachfragen zu überlasten.

In Krankenhäusern dürfen Entscheidungen aus medizinischen Gesichtspunkten nicht autonom von IS getroffen werden, sondern müssen von Menschen verantwortet werden. Aus diesem Grund können die Softwareagenten neue Behandlungstermine nur vorschlagen; die Termine müssen dann durch die menschlichen Akteure bestätigt oder verworfen werden. In der Konsequenz entsteht im Zusammenspiel von IS und Menschen die „hybride Intelligenz“ [Ram03]: Mensch und IS können Entscheidungen nicht alleine treffen, sondern bedingen sich gegenseitig.

4.2.3 Effektoren

Um die Entscheidungsvorschläge der Agenten wieder in die Wirklichkeit kommunizieren zu können, muss eine Schnittstelle zwischen Agenten und Menschen bestehen. Die Agenten benötigen Effektoren, um die Wirklichkeit zu beeinflussen. Dieser Kommunikationsweg soll über mobile Endgeräte realisiert werden. PDAs oder Handys bieten sich dabei an, da sie über bidirektionale Kommunikationswege verfügen. Die Agenten können darüber die Menschen über Entscheidungsvorschläge informieren, die Menschen können diesen Vorschlägen dann direkt zustimmen oder sie ablehnen. Die Agenten nutzen diese neue Information (Zustimmung oder Ablehnung), um die logische Welt entsprechen zu aktualisieren und die anderen Agenten über die Veränderung zu informieren. Im Sinne der ruhigen Technologie werten die Agenten auch ein Nichtbeachten eines Terminvorschlags innerhalb einer bestimmten Zeitspanne als Ablehnung. Zu Grunde liegt die Annahme, dass Ärzte nicht auf Meldungen reagieren, weil sie beschäftigt sind und dem PDA keine Aufmerksamkeit widmen können.

Die Softwareagenten optimieren den Behandlungsplan des Gesamtsystems Krankenhaus und können durch wechselseitige Verhandlungen auch kurzfristig auftretende Störungen im Terminplan beheben, indem die fehlende Ressource durch eine andere geeignete ersetzt wird. Welche Ressource aus dem Pool an verfügbaren Ressourcen ausgewählt wird, hängt vom Geldsubstitut ab, welches die Priorisierung der Ressourcen übernimmt.

5 Inhaltliches Vorgehen

Das Artefakt in der elementaren Funktionalität bildet die Grundlage der Arbeit. Die Grundfunktionalität eignet sich jedoch nicht, die von Weiser geforderte ruhige Technologie umzusetzen, da die menschlichen Akteure noch immer mit Entscheidungen behelligt werden, die von IS getroffen werden könnten. Das Artefakt wird daher in verschiedene Richtungen erweitert. Abbildung 2 stellt diese Erweiterungen grafisch dar.

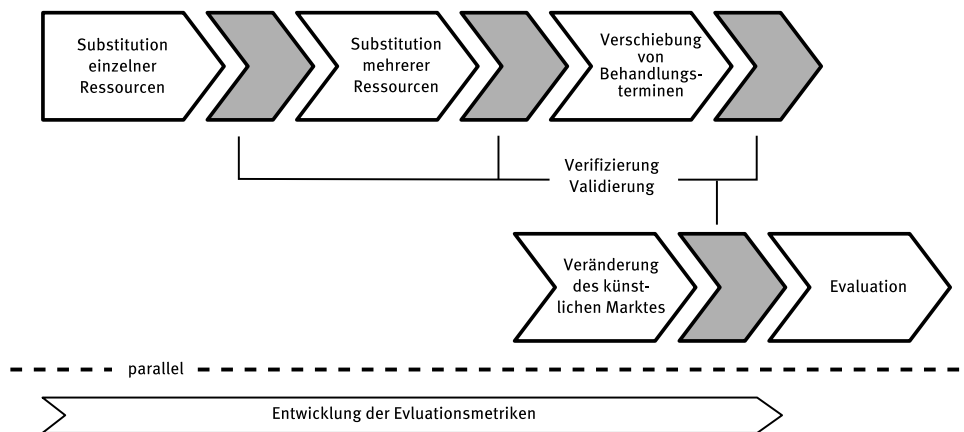


Abbildung 2: Einzelne Phasen der Dissertation. Die Entwicklung der Evaluationsmetriken erfolgt parallel zu den anderen Phasen. In der abschließenden Evaluationsphase kommen die Metriken zur Anwendung.

5.1 Substitution mehrerer Ressourcen

EMIKA kann nur die Ressourceklasse der Ärzte ersetzen, während alle anderen Ressourcen nur über die Verschiebungen informiert werden. Diese Einschränkung wird aufgehoben, so dass alle Anbieter-Ressourcen durch wechselseitige Verhandlungen ersetzt werden können. Die Problemstruktur ändert sich damit vom Allokationsproblem zum Koallokationsproblem, da nicht mehr nur eine Ressource auf dem künstlichen Markt eingekauft werden muss, sondern mehrere, die zum gleichen Zeitpunkt zur Verfügung stehen.

Mit dieser Erweiterung geht einher, dass sich Verhandlungen erheblich komplizierter gestalten: Neben dem Preis einer einzelnen Ressource müssen jetzt auch die Preise anderer benötigter Ressourcen miteinbezogen werden.

5.2 Verschiebung von Terminen

Da das bisherige Modell nur die Substitution von Ressourcen erlaubt, kann es vorkommen, dass ein Termin nicht stattfinden kann, weil nicht alle benötigten Ressourcen zur Verfügung stehen

und auch über Verhandlungen kein adäquater Ersatz gefunden werden kann. Solche Termine verfallen im bisherigen Modell. Diese Einschränkung wird aufgelöst, um absichern zu können, dass alle Behandlungstermine tatsächlich stattfinden. Dazu müssen Behandlungspfade hinterlegt werden, die die medizinischen Rahmenbedingungen abdecken, die die Terminvergabe beeinflussen (z. B. festgelegte Wartezeiten, oder die zwingende Folgebehandlung innerhalb eines bestimmten Zeitraums).

Das Einbeziehen von nachfolgenden Verhandlungen bzw. gesamter Behandlungspfade zieht eine Komplexitätssteigerung der zu verhandelnden Güter nach sich: Ein Agent kann nun nicht mehr sein gesamtes verfügbares Budget für den Einkauf einer Ressource einsetzen, sondern muss nachfolgende Behandlungen in seine Planung einbeziehen, da er auch für diese sein Budget verwenden muss. Auf dem Markt werden damit auch Güter mit einem zeitlichen Bezug gehandelt, während im bisherigen Modell nur Güter eingekauft werden, deren Nutzung sofort anfällt.

Mit der Erweiterung auf die Verschiebung von Terminen erweitert sich die Wissensbasis der Agenten: Sie müssen nun den gesamten Behandlungsplan in ihre Planung einbeziehen, während sie bisher nur die aktuelle Behandlung berücksichtigen mussten. Die Wissensbasis muss daher den gesamten klinischen Behandlungspfad des einzelnen Agenten umfassen.

5.3 Veränderung des künstlichen Marktes

Der Markt auf dem die Softwareagenten verhandeln, erfüllt nicht alle Bedingungen, um einem echten Markt zu entsprechen. Insbesondere ist die Budget- und die Preisfunktion nicht abhängig von den Verhandlungsergebnissen der Agenten. Wie viele Geldeinheiten eine Behandlung für einen Nachfrager kostet, hängt allein vom Beschäftigungszustand des Anbieters ab, nicht von den jeweiligen Marktpreisen. Desgleichen verfügt der EMIKA-Markt über einen offenen Geldkreislauf. Durch Wartezeit entsteht Geld, das durch den Einkauf von Ressourcen wieder abfließt. Vor diesem Hintergrund ist die Beurteilung von entstehenden Gleichgewichten schwierig. Deshalb wird der Marktmechanismus so angepasst, dass auch die Verhandlungsergebnisse in die Preis- bzw. Budgetfunktionen mit eingehen.

5.4 Evaluationsmetrik

Parallel wird eine Bewertungs-Metrik entwickelt. Um das Ziel der Effizienzsteigerung erreichen zu können, muss diese quantifizierbar sein. Aus den vorhandenen Metriken wird daher eine Aggregation entwickelt, die die Bewertung des Gesamtsystems erlaubt. Der aggregierte Nutzen

der einzelnen Softwareagenten bildet die „nutzenorientierte Wohlfahrt“ [Woo05, S. 106 ff.]. Neben der nutzenorientierten Wohlfahrt stehen weitere Nutzenmaße zur Verfügung zwischen denen abgewogen werden muss. Der fairnessorientierte Wohlfahrtsbegriff [Mou03] stellt durch eine anderen Aggregationsfunktion der Individualnutzen die Behandlung aller Akteure in den Vordergrund und sichert so ab, dass kein Patient nicht behandelt wird, weil sein geldwerter Nutzen zu niedrig ist. Die Auswahl des zu verwendenden Wohlfahrtsbegriffs findet im Rahmen der Entwicklung der Evaluationsmetrik statt.

6 Geplanter Abgabetermin

Die wissenschaftliche Arbeit und die Ausarbeitung sollte bis Ende 2008 abgeschlossen sein.

Literatur

- [BC06] *Becker, Marc; Czap, Hans*: Artificial Software Agents as Representatives of Their Human Principals in Operating-Room-Team-Forming. In: [KHLS06], S. 221–237.
- [CF03] *Clemens, Wolfgang; Fix, Walter*: Vom organischen Transistor zum Plastik- Chip. In: Physik Journal 2 (2003), Nr. 2, S. 31–36.
- [EMS06] *Eymann, Torsten; Müller, Günter; Strasser, Moritz*: Self-Organized Scheduling in Hospitals by Connection Agents and Mobile Devices. In: [KHLS06], S. 319–337.
- [Eym01] *Eymann, Torsten*: Markets without Makers – A Framework for Decentralized Economic Coordination in Multiagent Systems. In: *Fiege, Ludger; Mühl, Gero; Wilhelm, Uwe G. (Hrsg.): WELCOM Bd. 2232*, Springer, Berlin, Heidelberg 2001, S. 63–74.
- [Fin02] *Finkenzeller, Klaus*: RFID-Handbuch. 3. Aufl., Hanser, München 2002.
- [FYLA04] *Falck-Ytter, Yngve; Lang, Britta; Antes, Gerd*: Evidenzbasierte Medizin (EBM). In: *Jähn, Karl; Nagel, Eckhard (Hrsg.): e-Health*. Springer, Berlin 2004, Kapitel 3.1, S. 140–146.

- [HMPR04] *Hevner, Alan R.; March, Salvatore T.; Park, Jinsoo; Ram, Sudha*: Design Science in Information Systems Research. In: MIS Quarterly 28 (2004), März, Nr. 1, S. 75–105.
- [KHLS06] *Kirn, Stefan; Herzog, Otthein; Lokemann, Peter; Spaniol, Otto (Hrsg.)*: Multiagent Engineering – Theory and Applications in Enterprises. Springer, Berlin, Heidelberg 2006.
- [Kir06] *Kirn, Stefan*: Flexibility of Multiagent Systems. In: [KHLS06], S. 53–69.
- [Mae94] *Maes, Pattie*: Agents that Reduce Work and Information Overload. In: Communications of the ACM 37 (1994), Nr. 7, S. 30–40.
- [Mer95] *Mertens, Peter*: Wirtschaftsinformatik – Von den Moden zum Trend. In: *König, Wolfgang (Hrsg.)*: Wirtschaftsinformatik '95. Physica-Verlag, Heidelberg 1995, S. 25–64.
- [MG01] *Münch, Ines; Gnoth, Marcel*: ChariTime – Ein agentenorientiertes Softwaresystem zur Terminplanung im Krankenhaus. In: Wirtschaftsinformatik 43 (2001), Nr. 2, S. 183–187.
- [MKS+03] *Müller, Günter; Kreutzer, Michael; Strasser, Moritz; Eymann, Torsten; Hohl, Adolf; Nopper, Norbert; Sackmann, Stefan; Coroama, Vlad*: Geduldige Technologie für ungeduldige Patienten – Führt Ubiquitous Computing zu mehr Selbstbestimmung? In: *Mattern, Friedemann (Hrsg.)*: Total Vernetzt. Springer, Berlin, Heidelberg 2003, S. 159–186.
- [Mou03] *Moulin, Hervé*: Fair division and collective welfare. MIT Press, Cambridge, Massachusetts; London 2003.
- [NE06] *Niemann, Christoph; Eymann, Torsten*: Softwareagenten in der Krankenhauslogistik – ein Ansatz zur effizienten Ressourcenallokation. In: HMD: Praxis der Wirtschaftsinformatik 251 (2006), S. 77–87.

- [PZR+06] *Paulussen, Torsten O.; Zöller, Anja; Rothlauf, Franz; Heinzl, Armin; Braubach, Lars; Pokahr, Alexander; Lammersdorf, Winfried*: Agent-Based Patient Scheduling in Hospitals. In: [KHLS06], S. 255–275.
- [Ram03] *Rammert, Werner*: Die Zukunft der künstlichen Intelligenz: verkörpert – verteilt – hybrid / Technische Universität Berlin 2003 (TUTS-WP-4-2003). – working paper.
- [SE04] *Strasser, Moritz; Eymann, Torsten*: Selbstabstimmung von Terminplänen mit Hilfe zeitabhängiger Geldsubstitute. In: Mensch und Computer, 2004.
- [VK05] *Vaishnavi, Vijay; Kuechler, Bill*: Design Research in Information Systems / Association for Information Systems. 2005. – Forschungsbericht.
- [Wei91] *Weiser, Mark*: The Computer for the 21st Century. In: Scientific American 265 (1991), S. 66–75.
- [Wei98] *Weiser, Mark*: The Coming Age of Calm Technology. In: *Denning, Peter J.; Metcalfe, Robert M. (Hrsg.): Beyond Calculation: The Next Fifty Years of Computing*. Springer, Heidelberg 1998, Kapitel 6.
- [Wei96] *Wellman, Michael P.*: Market-Oriented Programming: Some Early Lessons. In: *Clearwater, Scott H. (Hrsg.): Market-Based Control*. World Scientific, 1996, Kapitel 4, S. 74–95.
- [Woo05] *Wooldridge, Michael J.*: An Introduction to MultiAgent Systems. John Wiley & Sons, Ltd 2005.

Kontextbasierte Adaption Klinischer Pfade

Claudia Reuter

Fraunhofer Institut für Software- und Systemtechnik (ISST)
Emil-Figge-Str. 91
44227 Dortmund
reuter@do.isst.fhg.de

Abstract

Klinische Pfade gewinnen im Krankenhaussektor als Mechanismus zur Steuerung von Behandlungsabläufen zunehmend an Bedeutung. Aktuelle Ansätze zur Realisierung der Pfade berücksichtigen die Anforderungen im Hinblick auf die Dynamik im medizinischen Umfeld nur unzureichend. Im Rahmen dieses Dissertationsvorhabens wird ein Konzept vorgestellt, dass die technische Umsetzung klinischer Pfade flexibilisiert und die Ausrichtung der Behandlung an patienten- und institutionsspezifischen Besonderheiten erleichtern soll. Die kontextbasierte Adaption soll sowohl ad hoc als auch evolutionäre Anpassungen klinischer Pfade unterstützen. Auf diese Weise soll die Entwicklung der Pfade enger mit den täglichen Arbeitsabläufen verknüpft werden, so dass sie letztlich das Produkt aus einer Zusammenstellung der „Best Practices“ eines Krankenhauses bilden. Bei der Realisierung dieses Konzepts werden Ansätze des adaptiven Workflow Managements zur dynamischen Modifikation von Prozessen berücksichtigt.

1 Einleitung

Das Gesundheitswesen sieht sich vor der Herausforderung, Kosten für die medizinische Versorgung zu reduzieren und gleichzeitig die Qualität der Behandlung zu erhöhen. Die Einschränkung von Varianzen in der Patientenversorgung wird als der effektivste Weg erachtet, um den Spagat zwischen Qualität und Wirtschaftlichkeit zu bewältigen [Chea00a]. Unter dem

Schlagwort „Clinical Pathways“ bzw. Klinische Pfade / Behandlungspfade wird daher der Versuch unternommen, die Geschäftsprozesse eines Krankenhauses, nämlich die Patientenbehandlung, wo möglich zu standardisieren und dabei zu optimieren. Darüber hinaus sollen klinische Pfade dem Anspruch genügen, Erkenntnisse aus der evidenzbasierten Medizin in den ärztlichen Alltag zu implementieren. Obwohl diese Entwicklung bereits in den neunziger Jahren begonnen hat, ist sie wegen der Komplexität dieser Aufgabe bisher nur in wenigen Krankenhäusern ernsthaft angekommen. Dennoch wird erwartet, dass aufgrund des Kostendrucks im Gesundheitswesen, des steigenden Qualitätsbewusstseins in der Bevölkerung und der Diskussion um die rechtliche Bedeutung medizinischer Leitlinien [Lipp04] die Relevanz klinischer Pfade als Mittel zur Prozessoptimierung zunehmen wird.

Wichtiger als die Reduktion von Varianzen und die Optimierung des Versorgungsprozesses ist jedoch die Ausrichtung der Behandlung an den spezifischen Bedürfnissen von Patientinnen und Patienten. Im optimalen Fall entspricht der klinische Pfad bereits zu dem Zeitpunkt, an dem er auf einen Patienten angewendet wird, dessen individuellen Erfordernissen. Ändert sich z.B. der Gesundheitszustand des Patienten oder treten Komplikationen auf, sollte die Möglichkeit bestehen, den Pfad entsprechend der geänderten Rahmenbedingungen zu modifizieren. Das Ziel dieses Dissertationsvorhabens besteht darin, auf der Basis klinischer Pfade spezialisierte Pfade zu generieren, die an das Profil eines bestimmten Patienten angepasst sind. Der adaptive Charakter dieser Pfade soll sicherstellen, dass Änderungen des Patientenprofils sowie Variationen in Bezug auf Arbeitsabläufe und Ressourcenzuordnung während der Behandlung adäquat berücksichtigt werden können und der Aufwand für das medizinische Personal möglichst gering ist.

2 Bisherige Ansätze zur Realisierung klinischer Pfade

Da klinische Pfade den Behandlungsverlauf maßgeschneidert gemäß den medizinischen und organisatorischen Praktiken in einem bestimmten Krankenhaus strukturieren, ist ihre Entwicklung mit hohem Zeit- und Kostenaufwand verbunden. Aus diesem Grund ist die erfolgreiche Implementierung des klinischen Pfades von entscheidender Bedeutung. Im Folgenden werden bisherige Ansätze zur Realisierung klinischer Pfade in Krankenhäusern vorgestellt.

2.1 Definition und Eigenschaften klinischer Pfade

Mit der Entwicklung und Umsetzung klinischer Pfade werden viele Erwartungen verknüpft. Die folgende Definition soll die Zielvorstellungen, die man mit klinischen Pfaden verbindet, verdeutlichen: „Ein klinischer Behandlungspfad ist der im Behandlungsteam selbst gefundene berufsgruppen- und institutionsübergreifende Konsens für die beste Durchführung der gesamten stationären Behandlung unter Wahrung festgelegter Behandlungsqualität sowie unter Berücksichtigung der notwendigen und verfügbaren Ressourcen, ebenso unter Festlegung der Aufgaben sowie der Durchführungs- und Ergebnisverantwortlichkeiten. Der klinische Behandlungspfad steuert den Behandlungsprozess; gleichzeitig ist er das behandlungsbegleitende Dokumentationsinstrument und erlaubt die Kommentierung von Normabweichungen zum Zwecke fortgesetzter Evaluation und Verbesserung.“ [Roed03]. Klinische Pfade sind demnach in erster Linie Steuerungsinstrumente, die den Ablauf der Patientenversorgung gemäß den Organisationsstrukturen der Einrichtungen, in denen sie zum Einsatz kommen, beschreiben. Um den qualitativen Anforderungen gerecht zu werden, soll bei der Entwicklung klinischer Pfade Wissen aus der evidenzbasierten Medizin berücksichtigt werden. Dabei handelt es sich um Kenntnisse auf dem aktuellen Stand der Wissenschaft, das auf medizinisch-ökonomisch vernünftigen Regeln und Grundsätzen basiert (z.B. medizinische Leitlinien).

Aus ökonomischer Sicht müssen Behandlungspfade rentabel sein. Angesichts des Aufwands bei der Standardisierung medizinischer Prozesse ist die Häufigkeit einer bestimmten Diagnose in Verbindung mit hohen Fallkosten eine wichtige Voraussetzung für die Erstellung eines klinischen Pfades, da hier die größten Einsparungspotentiale zu erwarten sind [Rahn02]. Darüber hinaus ist aufgrund mangelnder Flexibilität hinsichtlich der klinischen Parameter (z.B. Alter, Geschlecht, Stadium der Erkrankung, usw.) und des Ressourcenaufwands auf eine möglichst homogene Patientengruppe zu achten [Roed03]. Nur so kann sichergestellt werden, dass der Pfad häufig zur Anwendung kommt und sich seine Entwicklung rechnet.

Kommt es im Laufe der Behandlung zu einer Situation, die bei der Entwicklung des klinischen Pfades nicht berücksichtigt wurde, so ist eine Abweichung vom Pfad erforderlich. Die Gründe für ein solches Abweichen müssen entsprechend dokumentiert werden (Varianzdokumentation). Die Ergebnisse aus der späteren Varianzanalyse werden zur Evaluation und Pflege des klinischen Pfades verwendet.

2.2 Papierbasierte klinische Pfade

Aktuell werden klinische Pfade häufig in Papierform in Kliniken eingeführt. Da hier die Einhaltung des Pfades getrennt von den übrigen medizinischen Daten dokumentiert wird, führt diese Methode zu Medienbrüchen und zu Mehraufwand für ärztliches und pflegerisches Personal. So zog Cornelia Müller-Wenzel, Leiterin des zentralen Controlling im Knappschaftskrankenhaus Bottrop, nach der praktischen Erprobung der ersten papierbasierten klinischen Pfade den Schluss, deren Anwendung sei „sehr kompliziert und mit einem hohen Aufwand verbunden. Ärzte und Pflegekräfte mussten sowohl in den Pfad dokumentieren als auch die normale Leistungsdokumentation vornehmen – also doppelte Arbeit leisten. Daher war die Akzeptanz recht gering und die Durchdringung lag bei lediglich 30 Prozent.“ [Voss06]. Eine Anpassung des Pfades an das aktuelle medizinische Profil eines Patienten ist auf der Basis von Papierdokumenten nicht möglich. Erweist sich ein klinischer Pfad für die Versorgung eines bestimmten Patienten als unzureichend, ist der Grund für die Varianz vor dem Verlassen des klinischen Pfades zu dokumentieren. Hier wurde die Erfahrung gemacht, dass „die Abweichungsbögen teilweise nicht ausgefüllt und schwer lesbar waren“ [Voss06]. Auf diese Weise wird die Analyse der Varianzen erschwert, was sich wiederum negativ auf die Pflege und Weiterentwicklung klinischer Pfade auswirkt.

2.3 Integration klinischer Pfade in Primärsysteme

Aufgrund der Probleme bei einer papierbasierten Implementierung von Behandlungspfaden wird zunehmend die direkte Integration der Pfade in existierende Krankenhausinformationssysteme (KIS) angestrebt. Daher vollzieht sich im Krankenhausumfeld sowohl auf organisatorischer als auch auf technischer Ebene ein Übergang von einer vorwiegend funktions- und datenorientierten Dokumentation hin zu einer zunehmend bereichsübergreifenden, prozessorientierten Informationsverarbeitung. Die Prozessorientierung dieser Systeme wird in [Haas03] sogar als „das Paradigma der Zukunft“ bezeichnet, da nur so die Dokumentation medizinischer Leistungen entlang von Behandlungspfaden mit vertretbarem Aufwand ermöglicht werden kann.

KIS Hersteller verfolgen aktuell zwei Ansätze zur Entwicklung und Realisierung klinischer Pfade. Das eine Konzept sieht die Modellierung klinischer Pfade als Workflows vor. Um die Übersichtlichkeit des Pfades sicher zu stellen und die Wiederverwendung einzelner Pfadabschnitte zu ermöglichen, wird der Behandlungspfad häufig komposit aufgebaut. Die oberste Ebene ist z.B. entlang einer Zeitachse (präoperativer Tag, OP-Tag, erster postoperativer Tag, usw.) strukturiert. Bei den Aktivitäten innerhalb des Workflows handelt es sich demnach

entweder um weitere Pfade oder um elementare Aktionen, die von einem Verantwortlichen direkt ausgeführt werden können (z.B. Anamnesebogen ausfüllen). Der Workflow sieht sowohl sequentielle als auch parallele und alternative Zweige vor. Gibt es z.B. unterschiedliche Therapiemethoden für ein Krankheitsbild, so können die Parameter (Blutdruck, Risikofaktoren, usw.), die für die Entscheidung erforderlich sind, innerhalb eines speziellen Entscheidungsobjektes beschrieben werden. Bei der Ausführung des Pfades werden gemäß dem Pfadmodell Arbeitslisten für das ärztliche und pflegerische Personal generiert. Die Reihenfolge der Arbeiten orientiert sich dabei an den Vorgaben aus dem Modell. Vorteil dieses Lösungsansatzes ist die klare Ablaufstrukturierung; neben organisatorischen Aspekten können durch die Darstellung von Entscheidungsobjekten und alternativen Behandlungswegen ärztliche Entscheidungen explizit gemacht werden. Diese Ausdrucksfähigkeit bei der Modellierung erleichtert die Beschreibung aktueller Informationsbedarfe und damit die Integration von medizinischen Daten aus externen Wissensquellen, wie z.B. dem Patientenprofil oder medizinischen Leitlinien. Nachteilig wirkt sich hingegen die mangelnde Flexibilität der Lösung aus. Zwar können ggf. einzelne Schritte übersprungen werden, eine dynamische Anpassung des Pfades z.B. durch das Hinzufügen weiterer Behandlungsschritte ist jedoch nicht möglich. Im Falle einer Abweichung wird daher lediglich ein Dialog für die Abweichungsdokumentation geöffnet. Anschließend muss der Pfad verlassen werden.

Der zweite Ansatz modelliert klinische Pfade nicht als komplette Workflows, sondern nach dem Baukastenprinzip. Medizinische Aktivitäten, wie z.B. Röntgenbild anfordern oder Arztbrief schreiben stellen einzelne Bausteine dar, die beliebig angeordnet werden können. Für jeden Baustein kann der Start- und Endzeitpunkt sowie die zeitliche Dauer festgelegt werden. Klinische Pfade sind dabei nichts anderes als ein vordefiniertes Set von Bausteinen. Wird ein Patient auf einen Pfad gesetzt, so werden die entsprechenden Aktivitäten in seiner Patientenakte angelegt. Obwohl die Reihenfolge prinzipiell vorgegeben wird, kann der Arzt die Aktivitäten beliebig bearbeiten. Aus diesem Grund besteht der Vorteil dieser Lösung darin, dass Ärzte ein hohes Maß an Gestaltungsfreiheit besitzen. Wurde ein Pfad für einen Patienten initialisiert, können manuell Bausteine hinzugefügt oder entfernt werden, d.h. der klinische Pfad wird entsprechend den Erwägungen des Arztes variiert. Treten Abweichungen auf, wird der Arzt aufgefordert, diese Varianzen zu dokumentieren ohne den Pfad verlassen zu müssen. Allerdings ist die Ausdrucksfähigkeit dieser Art der Pfadmodellierung sehr begrenzt. So lässt sie z.B. keine Spezifikation von alternativen Behandlungszweigen zu. Außerdem können Entscheidungen, die

im Zuge der Patientenversorgung getroffen werden, nicht explizit dargestellt werden. Dieser Ansatz dient in erster Linie dazu, die medizinische Dokumentation und die Auftragsstellung an andere Abteilungen zu standardisieren. Aufgrund des geringen Grades an Ausdrucksfähigkeit bei der Pfadmodellierung sind die Möglichkeiten klinischer Pfade zur aktiven Steuerung des Behandlungsverlaufs begrenzt. Das Modell liefert daher wenig Informationen, die für eine Integration von Patientendaten zur Unterstützung bei der medizinischen Entscheidungsfindung und zur semi-automatischen Generierung eines patientenspezifischen Pfades benötigt werden.

Bisherige Realisierungen klinischer Pfade in Krankenhausinformationssystemen dienen in erster Linie der Strukturierung medizinischer Aktivitäten, der Reservierung benötigter Ressourcen und der Zuweisung von (Dokumentations-)Aufgaben an ärztliches und pflegerisches Personal. Weder das Workflow-Konzept noch das Baukastenprinzip zur Erstellung und Ausführung klinischer Pfade weist eine Schnittstelle zu einem Patientenprofil auf; eine direkte Einbeziehung von Patienteninformationen zur Adaption standardisierter Pfade ist daher nicht möglich. Auch welche Auswirkungen grundsätzliche Änderungen an einem Pfad im Rahmen der Varianzanalyse auf laufende Behandlungsprozesse entlang dieses Pfades haben, ist nicht geklärt.

3 Adaptives Workflow Management

Wie in Kapitel 2.1 beschrieben, stellen klinische Pfade ein Instrument zur Steuerung komplexer Behandlungsabläufe sowie zur Koordination der an der Patientenversorgung beteiligten Personen dar. Mit ihrer Hilfe soll die Transparenz von Entscheidungswegen erhöht werden. Aufgrund der engen Verknüpfung klinischer Pfade mit der evidenzbasierten Medizin ist ihr Einsatz auch zur Dokumentation einer qualitativ hochwertigen Behandlung (z.B. auf Basis medizinischer Leitlinien) mittelfristig absehbar. Aus diesem Grund ist es besonders wichtig, den Weg einer Behandlung sowie die Entscheidungen, die zu dieser Vorgehensweise geführt haben, nachvollziehen zu können. Die Workflow-Technologie stellt ausdrucksstarke Werkzeuge zur Modellierung und koordinierten Ausführung von Arbeitsprozessen, zur Ressourcenplanung, zur Organisation des Informationsflusses und zur Entscheidungsunterstützung bereit. Es liegt daher nahe, die zur expliziten Unterstützung von Geschäftsprozessen entwickelte Workflow-Technologie auf ihr Potential zur Abbildung komplexer, medizinischer Behandlungsabläufe bzw. Behandlungspfade zu untersuchen. Bereits jetzt kommen Workflow Management Systeme

(WfMS) in den klinischen Primärsystemen zum Einsatz, wie der in Abschnitt 2.3 beschriebene Ansatz zur Realisierung klinischer Pfade als Workflows zeigt.

Typische WfMS helfen bei der Definition, Ausführung, Koordination und Überwachung von Geschäftsprozessen innerhalb einer Organisation. Zu diesem Zweck muss die Struktur der Arbeitsprozesse und Aktivitäten durch eine maschinell interpretierbare Repräsentation, das Workflow-Modell, beschrieben werden. Der Modellierungsphase eines WfMS folgt dessen Laufzeitphase: Es reagiert auf vordefinierte Ereignisse mit der Instanziierung des passenden Workflow-Modells. Die vom WfMS verwaltete Instanz des Workflow-Modells stellt schließlich den konkreten ausführbaren Workflow dar [Deit97].

Traditionell waren Änderungen in WfMS nicht vorgesehen. Der Ablauf sollte im Voraus bekannt sein und in einer Prozessdefinition modelliert werden, die nahezu alle Fälle, die im Rahmen des Geschäftsprozesses bearbeitet werden, berücksichtigt. Auch an klinische Pfade stellt sich die Anforderung, zwischen 60 und 80 Prozent der Patientenbehandlungen abzudecken [Chea00b]. Dies ist eine Erklärung für die restriktive Auswahl von Behandlungssituationen, die mit Hilfe klinischer Pfade durchgeführt werden sollen, den hohen Zeit- und Kostenaufwand bei der Erstellung der Pfade und für Fehlschläge bei ihrer Implementierung. Beim Versuch, WfMS auch in Domänen einzusetzen, die eine hohe Komplexität und Dynamik aufweisen, hat sich die herkömmliche Vorgehensweise bei der Modellierung und Umsetzung als zu restriktiv herausgestellt [Hein99]. Basierend auf diesen Erkenntnissen wurde der Ansatz adaptiver WfMS entwickelt. Diese sollen, ähnlich wie herkömmliche Produktions-WfMS [Hast99], Geschäftsprozesse unterstützen, mit dem Unterschied, dass sie in der Lage sind, flexibel auf Änderungen zu reagieren. Diese Anforderung umfasst zwei fundamentale Aspekte [VdA99a]:

- Ad-hoc Adaptionen beziehen sich auf einzelne oder eine begrenzte Anzahl von Prozessen. Sie ermöglichen die spontane Reaktion auf geänderte Rahmenbedingungen in Bezug auf einen konkreten Fall.
- Bei evolutionären Adaptionen hingegen sind alle zukünftigen oder auch bereits laufenden Prozesse von der Änderung betroffen. Modifikationen an der Prozessstruktur sind Konsequenz einer grundsätzlichen Restrukturierung z.B. aufgrund korrekturbedürftiger, fehlerhafter Annahmen bei der Modellierung.

[Hein99] und [Dada00] beschreiben eine Reihe von Lösungsansätzen, um diese zwei Arten der Anpassung in WfMS zu realisieren. Mit Techniken des „Advance Modeling“ und „Late

Modeling“ / „Late Binding“ wird versucht, die notwendigen Modifikationen auf der Ebene des Workflow Modells vorzunehmen. Während bei „Advance Modeling“ alle möglichen Alternativen bereits während der Prozessdefinition berücksichtigt werden müssen, erlaubt „Late Modeling“ / „Late Binding“ das Auslassen bestimmter Abschnitte, um das Workflow Modell nicht mit Informationen zu überfrachten. Die Auswahl bzw. Modellierung der konkreten Prozessschritte, die diese offenen Abschnitte, so genannte „Black Boxes“ füllen, erfolgt zur Laufzeit des Prozesses. Diesen Ansatz verfolgt z.B. das modulare WfMS „MOBILE“ [Hein99]. Es betrachtet Workflows aus unterschiedlichen Perspektiven, wie z.B. die Verhaltensperspektive, die den Kontrollfluss betrachtet, oder die Informationsperspektive, die sich auf den Datenfluss bezieht. Mit Hilfe der MOBILE Modeling language (MOMO) können einzelne Perspektiven innerhalb des Workflow Modells unspezifiziert bleiben. Ist z.B. zum Zeitpunkt der Prozessentwicklung nicht klar, in welcher Reihenfolge eine Anzahl von Aktivitäten durchzuführen ist, wird dieser Teil der Prozessdefinition offen gelassen und erst während der Prozessausführung festgelegt. Der Nachteil dieses Lösungsansatzes besteht darin, dass im Voraus bekannt sein muss, was spontan während der Laufzeit des Prozesses zu entscheiden ist und an welcher Stelle im Workflow Modell Auslassungen erfolgen müssen. Ad hoc und evolutionäre Anpassungen können daher nur in begrenztem Ausmaß unterstützt werden.

Eine weitere Methode wird als „Ad Hoc Adaption“ / „Instance Adaption“ bezeichnet. Dieser Ansatz folgt der Annahme, dass Anpassungen am Prozess zu jedem Zeitpunkt der Ausführung notwendig werden können und eine Festlegung des Auftretens dieser Änderungen a priori unmöglich ist. Eine große Herausforderung der „Ad Hoc Adaption“ ist die Sicherstellung der Korrektheit eines Workflows nach der Anpassung. Außerdem ist die Frage zu klären, welche Modifikationen (semi-)automatisch erfolgen können und was ausschließlich manuell vorgenommen werden kann. Die Arbeitsgruppe in Ulm um Dadam und Reichert greift die Problematik der ad hoc Anpassung auf und beschreibt mit ADEPT (Application Development based on Encapsulated Pre-modeled activity Templates) einen Ansatz, der es ermöglicht, planbare Abweichungen vom Standardablauf eines Arbeitsprozesses bereits zur Modellierzeit festzulegen und die Durchführung während der Laufzeit des Prozesses semi-automatisch zu unterstützen [Reic00][Reic02]. Ein grundsätzliches Problem bei der Modellierung von geplanten Abweichungen besteht darin, dass meist nicht alle Varianzen modelliert werden

können. Darüber hinaus berücksichtigt der Ansatz nicht die Beziehung geänderter Workflow Instanzen zum ursprünglichen Prozessmodell.

Ansätze der Kategorie „Schema Evolution“ / „Workflow Definition Adaption“ nehmen Anpassungen nicht an einzelnen Prozessinstanzen vor, sondern modifizieren das zugrunde liegende Workflow Modell; anschließend werden diese Änderungen an alle zukünftigen oder auch bereits laufenden Instanzen propagiert. Während die „Ad hoc Adaption“ in erster Linie spontane Anpassungen von Prozessen fokussiert, adressiert diese Methode vor allem die Evolution von Geschäftsprozessen. Dieses Ziel verfolgt Van der Aalst [VdA99b] mit seinem Ansatz des „generischen Workflow Modells“. Es beschreibt Klassen von Varianten desselben Workflows. Inspiriert ist dieser Lösungsansatz durch Techniken der industriellen „Product Configuration“ [Veen92]. Damit wurde auf die Anforderung reagiert, dasselbe Produkt in verschiedenen Ausfertigungen zur Verfügung zu stellen. Van der Aalst erweitert das klassische Workflow Modell um eine horizontale Navigationskomponente, bei der es sich um die Generalisierung bzw. Spezialisierung von Modellen handelt. Im Gegensatz dazu bezieht sich die vertikale Navigationskomponente auf die Komposition von Workflow Abschnitten zu einem Gesamtprozess. Die Generalisierung von Prozessen erfordert die Spezifikation eines generischen Workflows, auf den alle übrigen Workflows abgebildet werden können. Die Spezialisierung erfolgt durch eine Umkehrung dieser Abbildung. Auf diese Weise soll das Problem dynamischer Änderungen von Prozessen auf definierte Zustandsübergänge reduziert werden. Eine Schwäche dieses Ansatzes besteht daher in einem sehr aufwändigen Modellierungsprozess der einzelnen Varianten und ihren Übergängen, sowie dem Informationsverlust, der beim Wechsel von einer komplexen zu einer einfacheren Prozessdefinition auftritt.

Adaptive WfMS sollen die Vorteile herkömmlicher WfMS nutzen und zugleich flexiblere Konzepte bereitstellen, die die besonderen Erfordernisse dynamischer Arbeitsumgebungen berücksichtigen. Die vorgestellten Ansätze zur Modellierung und Umsetzung adaptiver WfMS adressieren jeweils unterschiedliche Aspekte dieser komplexen Anforderung. Einige Lösungen setzen den Fokus auf eine detaillierte Prozessentwicklung und -beschreibung, andere konzentrieren sich auf dynamische Änderungen von Prozessinstanzen. Im Hinblick auf die äußerst komplexen, umfangreichen und qualitätsgesicherten Behandlungsprozesse im Gesundheitswesen sind alle Aspekte von Bedeutung. Einerseits ist die detaillierte Strukturierung der Soll-Prozesse im Krankenhaus notwendig, andererseits ist die medizinische

Domäne zu vielschichtig, um bei der Entwicklung klinischer Pfade alle alternativen Behandlungswege im Voraus berücksichtigen zu können. Im Laufe der Behandlung kommt es häufig zu ad hoc Entscheidungen und neben planbaren Vorkommnissen treten auch unvorhergesehene Ereignisse auf. Klinische Pfade dürfen demnach nicht ausschließlich Produkt eines aufwändigen Abstimmungsprozesses zwischen Fachleuten unterschiedlicher Disziplinen sein, sondern sollten sich flexibel an die Rahmenbedingungen ihres aktuellen Anwendungskontexts anpassen lassen.

4 Kontextbasierte Adaption klinischer Pfade

Bei der Diskussion der bisherigen Ansätze zur Realisierung klinischer Pfade und bei der Darstellung von Konzepten des adaptiven Workflow Managements, hat sich gezeigt, dass die vorhandenen Methoden für eine adäquate Umsetzung von Behandlungspfaden nicht ausreichen. Klinische Pfade sind keine starren Konstrukte, sondern müssen stets die individuellen Anforderungen eines Patienten berücksichtigen, kontinuierlich revidiert und an bewährte Praktiken angepasst werden. Was aktuell bei der Implementierung klinischer Pfade unzureichend berücksichtigt wird, ist die Diskrepanz zwischen geplanter und tatsächlicher Arbeit. Bei der Planung eines klinischen Pfades ist es nicht möglich, den konkreten Behandlungskontext, in dem der Pfad später zur Anwendung kommt, exakt vorherzusehen. Dies wird zwar durch die Definition einer möglichst homogenen Patientengruppe versucht, aber selbst wenn ein Patient anfangs den Parametern dieser Gruppe entspricht, heißt das nicht, dass dies während des gesamten Behandlungsverlaufs auch der Fall ist. Klinische Pfade können daher als Arbeitspläne betrachtet werden, die ein gemeinsames Ziel definieren und den effizientesten Weg zur Erreichung dieses Ziels beschreiben. Da die konstante Einhaltung dieser Vorgaben im Laufe der Behandlung nicht immer möglich ist, der klinische Pfad als Mittel zur Koordination der Patientenversorgung jedoch beibehalten werden sollte, müssen sich Pfade flexibel an aktuelle Arbeitsbedingungen anpassen lassen. Diese Fähigkeit bezeichnet [Bard97] auch als „situated planning“. Der Ansatz des „situated planning“ stützt sich auf Erkenntnisse der Aktivitätstheorie, die einen Plan folgendermaßen definiert: „Cognitive or material artifact which supports the anticipatory reflection of future goals for actions, based on experience about recurrent structures in life“ [Bard97]. Die Übertragung eines Plans von der Theorie in die Praxis

erfordert seine Anpassung an die konkreten Arbeitsbedingungen und führt zu einem Lernprozess, der die Weiterentwicklung und Verbesserung des Plans ermöglicht.

4.1 Anforderungen

Betrachtet man klinische Pfade unter dem Aspekt des „situated planning“, lassen sich neben den in 2.1 spezifizierten Eigenschaften folgende zusätzliche Anforderungen definieren:

- Klinische Pfade sollten immer als Teil des Versorgungsprozesses, in dem sie zum Einsatz kommen, betrachtet werden. Aus diesem Grund ist es notwendig, die kontinuierliche Entwicklung und Modifikation eines Behandlungspfades basierend auf den Erfahrungen seiner praktischen Ausführung zu ermöglichen.
- Die Realisierung adaptiver klinischer Pfade muss berücksichtigen, dass in der konkreten Arbeitsumgebung, in der ein Pfad eingesetzt wird, andere Rahmenbedingungen gelten, als bei der Pfadentwicklung angenommen. Da das Pfadmodell ausschließlich den Behandlungsweg für einen Patienten unter optimalen Bedingungen beschreibt, muss die Möglichkeit bestehen, Pfade während ihrer Ausführung dynamisch anzupassen.

Die erste Anforderung kann zunächst mit Hilfe der Varianzdokumentation erfüllt werden. Diese ermöglicht es Ärzten, unter Angabe einer Begründung vom Pfad abzuweichen. Auf Basis der späteren Varianzanalyse kann festgestellt werden, ob das entwickelte Pfadmodell den Anforderungen der Praxis entspricht oder angepasst werden muss. Nachdem das Pfadmodell überarbeitet wurde, können alle zukünftigen Fälle gemäß diesem neuen Modell behandelt werden. Es kann jedoch sein, dass die Änderungen auch für diejenigen Patienten von Bedeutung sind, die auf eine alte Version des klinischen Pfades gesetzt wurden. Das ist z.B. dann der Fall, wenn sich bei der Pfadmodellierung Fehler eingeschlichen haben oder prinzipielle Arbeitsabläufe in der Klinik geändert wurden und es nicht tolerierbar ist, dass Patienten noch nach dem alten Muster behandelt werden. Damit in solchen Fällen keine manuellen Änderungen an einzelnen Pfaden notwendig sind, müssen Möglichkeiten bereit stehen, Änderungen am Pfadmodell für alle Patienten zu berücksichtigen, die auf den Pfad gesetzt wurden. Diese Anforderung wird auch im Rahmen des Konzepts adaptiver Workflows von [VdA99a] als evolutionäre Anpassung bezeichnet (siehe Kapitel 3).

Da klinische Pfade einen möglichst optimalen Behandlungsverlauf während eines stationären Aufenthalts beschreiben, ist es gar nicht Intention der Pfadentwickler, im Pfadmodell alle

Eventualitäten der Patientenversorgung zu berücksichtigen. Neben der Tatsache, dass dies nicht realisierbar ist, entspricht es auch nicht dem Wunsch des Krankenhausmanagements, sämtliche Behandlungswege über den klinischen Pfad abzubilden. Ist eine Einhaltung des klinischen Pfades bzw. des optimalen Behandlungsablaufs nicht möglich, so muss dies als Abweichung feststellbar sein. Dies gilt selbst dann, wenn die Maßnahmen zum Umgang mit einer Varianz, wie z.B. einer Komplikation, gut zu strukturieren sind. Damit unterscheidet sich die Realisierung adaptiver klinischer Pfade von der Umsetzung anderer (adaptiver) Geschäftsprozesse, bei denen eine Adaption nur deswegen erfolgen muss, weil es nicht möglich ist, ihren Verlauf komplett vorherzusagen. Da es während der Patientenversorgung jedoch immer wieder zu Ereignissen und Entscheidungen kommt, die eine Abweichung vom klinischen Pfad erforderlich machen, muss die Möglichkeit bestehen, den Pfad an die geänderten Bedingungen anzupassen. Bezüglich adaptiver Workflows spricht man bei dieser Art der Prozessanpassung von „Ad-hoc Adaption“ (siehe Kapitel 3). Bei den Vorkommnissen, die zu der Varianz führen, kann es sich dabei um planbare oder auch unvorhergesehene Ereignisse handeln. Zu den planbaren Ereignissen zählen z.B. bekannte Komplikationen oder zusätzliche Maßnahmen aufgrund einer bestimmten Vorerkrankung des Patienten. Einerseits muss sichergestellt werden, dass es sich dabei um Abweichungen von der optimalen Versorgung handelt; andererseits sollte es im Rahmen solcher planbaren Varianzen nicht notwendig sein, dieselben Änderungen jedes Mal neu vornehmen zu müssen. Es ist also wünschenswert, Anpassungen aufgrund eines bestimmten Ereignisses, die einmal erfasst und im Rahmen der Varianzanalyse validiert wurden, auf ähnliche Situationen anwenden zu können. Konsequenz ist, dass man im Rahmen klinischer Pfade zwischen Pfadelementen unterscheiden muss, mit denen der optimale Behandlungsweg beschrieben wird, und solchen, die planbare Abweichungen vorstrukturieren.

4.2 Zukunftsszenario

Der Nutzen einer kontextbasierten Adaption klinischer Pfade soll an Hand eines Zukunftsszenarios, das unterschiedliche Aspekte im Hinblick auf notwendige Anpassungen adressiert, verdeutlicht werden. Dem Szenario liegt ein papierbasierter klinischer Pfad des Knappschaftskrankenhauses Dortmund zur radikalen retropubischen Prostatektomie zugrunde. Dabei handelt es sich um das Standardverfahren zur kurativen Therapie des Prostatakarzinoms in einem frühen Stadium. Es beinhaltet die radikale Entfernung der Prostata mit teilweiser Entfernung der Lymphknoten im Becken.

Herr Müller ist 61 Jahre alt und wird wegen akuter Darmkolik in die Notaufnahme des Krankenhauses gebracht. Aufgrund seines Alters wird standardmäßig eine Krebsvorsorgeuntersuchung durchgeführt. Dabei ergeben sich erste Hinweise auf das Vorliegen eines Prostatakarzinoms. Weitere diagnostische Maßnahmen bestätigen diese Verdachtsdiagnose. Der Patient entscheidet sich für die kurative Therapie in Form einer radikalen retropubischen Prostatektomie. Für diese Therapiemethode wurde ein klinischer Pfad entwickelt, so dass die Behandlung mit Hilfe eines „Pfadplaner“-Systems gesteuert und dokumentiert werden kann. Bei der Identifikation der für Herrn Müller relevanten Behandlungspfade wird sowohl der Pfad „Radikale retropubische Prostatektomie“ als auch „Akute Darmkolik“ ausgewählt. Der Pfadplaner muss nun beide Pfade miteinander verschmelzen, so dass z.B. pflegerische Tätigkeiten besser koordiniert werden können und auch die Medikation entsprechend eingestellt wird.

Der elektronischen Patientenakte kann entnommen werden, dass der Patient zudem seit Jahren unter Diabetes mellitus Typ II leidet. Der klinische Pfad zur Prostatektomie sieht für solche Fälle optionale Leistungen vor, z.B. zusätzliche Laboruntersuchungen im Hinblick auf den Blutzucker-Tagesspiegel und Diabetiker-Kost, um auch eine optimale Behandlung von Diabetikern entlang des klinischen Pfades gewährleisten zu können. Sobald der Patient auf den Pfad gesetzt wird, prüft der Pfadplaner das Patientenprofil und schlägt vor, diese optionalen Leistungen in die Behandlung mit einzubeziehen. Die Bestätigung des verantwortlichen Arztes genügt, um die vordefinierten Anpassungen automatisch vorzunehmen und die Abweichung zu dokumentieren.

Nach der Aufnahme des Patienten wird der klinische Pfad überarbeitet. Dabei wird der Medikationsplan und die Art der Präparate geändert. Diese Änderung soll sowohl bei neuen als auch bei alten Patienten, bei denen die Medikation noch nicht angepasst wurde, wirksam werden. Auch der Pfad von Herrn Müller wird unter Berücksichtigung der Ressourcenplanung und nach Einwilligung des behandelnden Arztes automatisch angepasst.

Während der Operation des Patienten muss sehr viel Lymphmaterial entnommen werden (pelvine Lymphadenektomie). Da das Standardverfahren diesen Vorgang nicht vorsieht, muss die Pfadanpassung ad hoc erfolgen. Die notwendigen Anpassungsaktionen sind noch nicht im System hinterlegt und müssen daher manuell vorgenommen werden. Im Rahmen der späteren Varianzanalyse wird der Pfad von Herrn Müller mit Blick auf das Behandlungsergebnis betrachtet und schließlich als Varianz in den Pfadplaner aufgenommen. Wird eine

Lymphadenektomie in Zukunft bei anderen Patienten erforderlich, sind die notwendigen Modifikationen bekannt und können automatisch in den jeweiligen Pfad übernommen werden.

4.3 Methodik

Ziel der kontextbasierten Adaption klinischer Pfade ist die Generierung und Modifikation von Pfaden basierend auf Kontextinformationen. Diese Informationen lassen sich drei Kategorien zuordnen:

- Medizinische Kontextinformationen beziehen sich auf das individuelle Patientenprofil, indem z.B. Dauerdiagnosen, aber auch aktuelle Laborwerte hinterlegt sind.
- Organisatorische Kontextinformationen werden durch institutionsspezifische Arbeitsabläufe und vorhandene Ressourcen bestimmt.
- Soziale Kontextinformationen definieren die Rahmenbedingungen der Patientenversorgung; dazu zählt, dass ein Patient seine Einwilligung zu einer bestimmten Behandlungsform erteilen muss.

Das Konzept der kontextbasierten Adaption klinischer Pfade soll durch Abbildung 1 veranschaulicht werden.

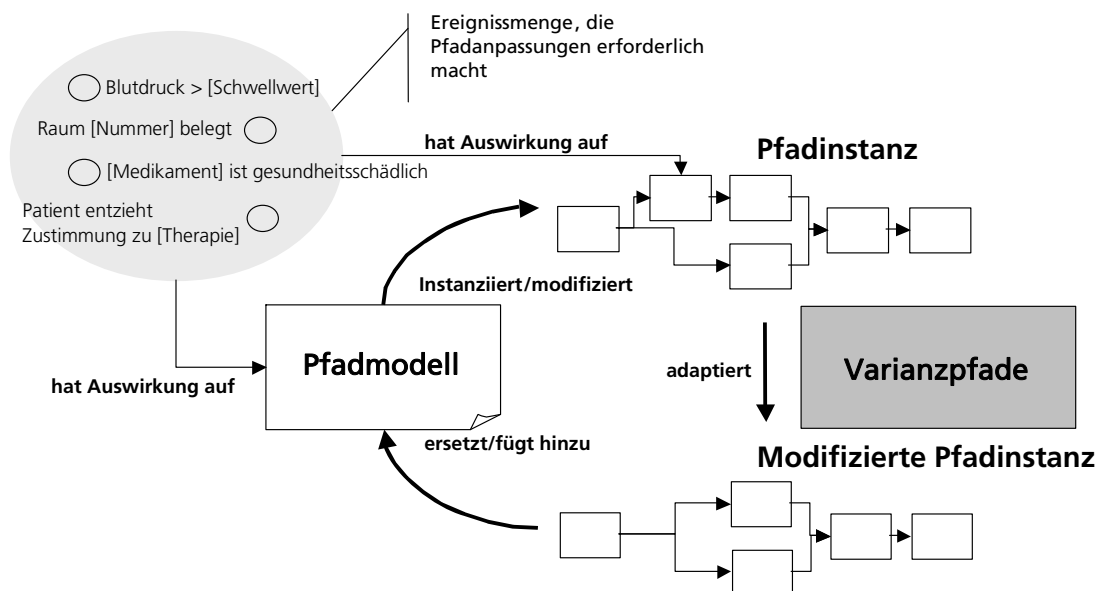


Abbildung 1 Kontextbasierte Adaption klinischer Pfade

Das Pfadmodell beschreibt den als optimal festgelegten Standardpfad; daneben existieren diverse Varianzen, die das Modell ergänzen bzw. in die das Modell überführt werden kann. Diese werden im Folgenden als „Varianzpfade“ bezeichnet. Wird ein Patient auf einen klinischen Pfad gesetzt, so wird für ihn analog zur Prozessinstanz von WfMS eine eigene Pfadinstanz erzeugt. Ob diese Instanz auf dem ursprünglichen Pfadmodell basiert oder bereits zu Beginn des Pfades Varianzen berücksichtigt werden müssen, hängt von den individuellen Merkmalen des Patienten (z.B. besondere Risiken, Vorerkrankungen, usw.) ab. Wie in Abbildung 1 dargestellt, können medizinische Kontextinformationen sowohl Auswirkungen auf das Pfadmodell, als auch auf alle oder einzelne Pfadinstanzen haben. Wird ein klinischer Pfad im Rahmen der Varianzanalyse bearbeitet oder stellt sich ein im Pfad referenziertes Medikament als allgemein gesundheitsschädlich heraus, so wird das Pfadmodell modifiziert. Sollen diese Änderungen für alle laufenden Behandlungsfälle wirksam werden, so ist es notwendig, auch die Pfadinstanzen an das geänderte Modell anzupassen. Dieser Fall wird im Hinblick auf adaptives Workflow Management auch als „Schema evaluation“ bezeichnet (siehe Kapitel 3).

Andere Kontextinformationen betreffen hingegen ausschließlich die Ausführung einzelner Pfadinstanzen. Steigt z.B. der Blutdruck des Patienten über einen tolerierbaren Schwellenwert, so müssen zunächst die betroffenen Aktivitäten innerhalb der Pfadinstanz identifiziert werden. Anschließend kann manuell oder gemäß bereits bekannter Varianzpfade die Anpassung der Pfadinstanz ad hoc erfolgen. Informationen über die geänderte Pfadinstanz müssen als Varianzdokumentation im System hinterlegt werden. Wird der Grund für die notwendige Varianz, z.B. mit Hilfe eines Patientenprofils das auf einen vom Sollwert abweichenden Laborwert verweist, automatisch erfasst, so kann diese Information auch in der nötigen Varianzdokumentation automatisch vermerkt werden.

Abweichungen auf Basis medizinischer Kontextinformationen können im Zuge der Varianzanalyse geprüft und ggf. dem ursprünglichen Pfadmodell als Varianzpfade hinzugefügt werden. Kommt es in einem anderen Behandlungsfall zu einer ähnlichen Konstellation, ist es möglich, unter Nutzung des bekannten Varianzpfades vom Standardverlauf abzuweichen, ohne den qualitätsgesicherten Versorgungsprozess verlassen zu müssen. Indem das medizinische Personal auf bekannte Verfahren bei der Behandlung von Varianzen zurückgreift, wird der Aufwand bei der Anpassung des Pfades an patientenspezifische Erfordernisse auf ein Minimum reduziert.

Organisatorische Kontextinformationen können sich wie der medizinische Kontext auf das Pfadmodell beziehen, z.B. wenn sich Arbeitsabläufe aufgrund struktureller Neuordnungen grundlegend ändern. Ebenso ist es möglich, dass einzelne Pfadinstanzen betroffen sind, z.B. wenn in Folge eines Notfalls ein OP-Raum belegt wird, der für Patienten vorgesehen war, die gemäß klinischem Pfad behandelt werden. Im Gegensatz zu medizinischen Kontextinformationen ist die Wahrscheinlichkeit, dass gleich mehrere Instanzen von der Änderung betroffen sind, bei organisatorischen Kontextinformationen höher. Soziale Kontextbedingungen können trotz Flexibilität zum Abbruch des Pfades führen; ist z.B. die Zustimmung des Patienten zur medizinischen Behandlung nicht länger gegeben, so muss auch der klinische Pfad beendet werden.

4.4 Werkzeuge

Diese Arbeit setzt den Schwerpunkt auf medizinische Kontextinformationen. Diese Informationen stammen in erster Linie von ärztlichen und pflegerischen Mitarbeitern. Aus diesem Grund muss die Möglichkeit bestehen, manuelle ad hoc Anpassungen am Pfad vorzunehmen. Im Hinblick auf die Workflow Management Technologie existieren bereits entsprechende Ansätze, wie z.B. ADEPT, das in Kapitel 3 skizziert wurde. Da die Fülle an medizinischen Informationen, besonders bei chronisch kranken Patienten, immer weiter ansteigt, sollten Ärzte bei der Nutzung dieser Daten unterstützt werden. In den vorhandenen Krankenhausinformationssystemen (KIS) gibt es oftmals bereits so genannte elektronische Patientenakten, in denen alle Informationen zu einem Patienten gespeichert sind. Diese Akten können daher ebenfalls als Quellen für medizinische Kontextinformationen dienen. Elektronische Patientenakten werden von den unterschiedlichen KIS jedoch sehr individuell angelegt und verwaltet. Aus diesem Grund ist es erforderlich, ein allgemeines Patientenprofil zu definieren. Dieses Profil sollte das Sammeln, Aggregieren und Darstellen von Patienteninformationen aus verschiedenen elektronischen Akten ermöglichen. Es umfasst Daten wie Dauerdiagnosen, Laborwerte, usw. die für die Behandlung eines Patienten entlang eines klinischen Pfades von Bedeutung sind. Für die Erstellung dieses Profils kann ein Standard, nämlich die XML-basierte Clinical Document Architecture (CDA) genutzt werden, die eine Verknüpfung mit HL7 zum Zweck des Sammelns relevanter Informationen erlaubt. Unter Beachtung der semantischen Interoperabilität sollte ein medizinischer Thesaurus zur Anwendung kommen. Um Pfadanpassungen durch die Präsentation bekannter Pfadvarianzen teilweise automatisieren zu können, müssen Regeln definiert werden, die auf aktuelle

Kontextinformationen im Patientenprofil angewendet werden und spezifizieren, welche Anpassungen vorzunehmen sind bzw. welche Varianzpfade bei diesem Behandlungsfall zum Tragen kommen. Grundlage für die Durchführung von Pfadanpassungen ist das adaptive Workflow Management.

5 Zusammenfassung

Kostendruck und gestiegenes Qualitätsbewusstsein im Gesundheitswesen haben dazu geführt, dass die Prozessabläufe in Krankenhäusern im Hinblick auf ihre Effizienz und Effektivität analysiert und kritisch diskutiert werden. Klinische Pfade werden zunehmend als Instrument zur Steuerung und Koordination der Behandlungsvorgänge eingesetzt. Bisherige Realisierungen werden dem Anspruch klinischer Pfade als flexible Handlungskorridore jedoch nur in unzureichendem Maße gerecht. Die meisten Ansätze zielen auf die Erleichterung von Dokumentationstätigkeiten ab oder definieren einen Versorgungsprozess ohne die Dynamik einer medizinischen Arbeitsumgebung angemessen zu berücksichtigen. Das adaptive Workflow Management adressiert die Anforderungen solcher dynamischen Umgebungen, indem es Möglichkeiten zur Flexibilisierung von Geschäftsprozessen bereitstellt. Bisherige Lösungsansätze beziehen sich jedoch nur auf Teilaspekte, die im Hinblick auf klinische Pfade von Bedeutung sind. Um die Diskrepanz zwischen Behandlungsplanung und –ausführung angemessen zu berücksichtigen, soll im Rahmen dieses Dissertationsvorhabens ein umfassendes Konzept zur kontextbasierten Adaption klinischer Pfade entwickelt werden. Auf diese Weise ist es möglich, Behandlungspfade an aktuell geltende Rahmenbedingungen im Hinblick auf individuelle Patientenanforderungen und dynamische klinische Arbeitsprozesse anzupassen. Zu diesem Zweck müssen neben dem als optimal definierten Behandlungsverlauf auch Varianzen zugelassen werden. Im Gegensatz zu Abweichungen von Geschäftsprozessen in anderen Domänen stellen Abweichungen von klinischen Pfaden keine Ausnahmesituationen dar, sondern gehören zum normalen klinischen Alltag und entsprechen der Notwendigkeit, die Versorgung an den spezifischen Bedürfnissen eines Patienten auszurichten. Ärzte sollten daher beim Erkennen einer Abweichung und bei ihrer Integration in den standardisierten klinischen Pfad entsprechend unterstützt werden. Neben ad hoc Anpassungen berücksichtigt das Konzept der kontextbasierten Adaption auch die evolutionäre Weiterentwicklung klinischer Pfade basierend auf den Erfahrungen während ihrer praktischen Anwendung.

Literaturverzeichnis

- [Bard97] Bardram JE: Plans as Situated Action: An Activity Theory Approach to Workflow Systems, Proc. of the Fifth European Conference on Computer Supported Cooperative Work, 1997, 17-32
- [Chea00a] Cheah J: Development and implementation of a clinical pathway programme in an acute care general hospital in Singapore, In: International Journal for Quality in Health Care, 2000, 12(5): 403-412
- [Chea00b] Cheah J: Clinical Pathways – An Evaluation of its Impact on the Quality of Care in an Acute Care General Hospital in Singapore, In: Singapore Med J, 2000, 41(7): 335-346
- [Dada00] Dadam P, Reichert M, Kuhn K: Clinical Workflows – The Killer Application for Process-oriented Information Systems? Proc. 4th Int. Conf. On Business Information Systems (BIS'2000), Poznan, Poland, Springer, Berlin, Germany: 36-59
- [Deit97] Deiters W: Prozessmodelle als Grundlage für ein systematisches Management von Geschäftsprozessen, In: Informatik Forschung und Entwicklung. Springer-Verlag, 1997, 12: 52-60
- [Haas03] Haas P: Aspekte handlungsunterstützender klinischer Informationssysteme, In: Klinikarzt, 2003, 32(11)
- [Hast99] Hastedt-Marckwardt, C: Workflow-Management-Systeme, In: Informatik-Spektrum, 1999, 22(2): 106
- [Hein99] Heintz P, Horn S, Jablonski S, Neeb J, Stein K, Teschke M: A Comprehensive Approach to Flexibility in Workflow Management Systems, ACM SIGSOFT Software Engineering Notes, 1999, 24(2): 79-88
- [Lipp04] Lippert HD: Leitlinien – rechtlich gesehen, In: Notfall & Rettungsmedizin, Springer Berlin / Heidelberg, 2004, 5(6): 451-453

- [Rahn02] Rahn A, Schena R: Tagungsbericht Klinische Pfade – Wohin führen sie?, http://www.kliniken-heh.de/Download_dat/Tagungsbericht_Klinische_Pfade_HEH_01.06.02.pdf, abgerufen am 28.11.2006
- [Reic00] Reichert M: Dynamische Ablaufänderungen in Workflow-Management-Systemen, Doktorarbeit, Universität Ulm, Mai 2000
- [Reic02] Reichert M, Bauer T, Fries T, Dadam P: Modellierung planbarer Abweichungen in Workflow Management Systemen, Modellierung in der Praxis – Modellierung für die Praxis, LNI, 2002, 12: 183-194
- [Roed03] Roeder N, Hindle D, Loskamp N, Juhra C, Hensen P, Rochell B: Frischer Wind mit klinischen Behandlungspfaden (I+II). Instrumente zur Verbesserung der Organisation klinischer Prozesse, 2003, das Krankenhaus, 1:20-27, 2:124-130
- [VdA99a] Van der Aalst WMP, Basten T, Verbeek HWM, Verkoulen PAC, Voorhoeve M: Adaptive Workflow: On the interplay between flexibility and support, 1998, In Filipe J, Cordeiro J (eds.), Proc. of the first Int. Conf. On Enterprise Information Systems, 1999, 353-360
- [VdA99b] Van der Aalst WMP: Generic Workflow Models: How to Handle Dynamic Change and Capture Management Information, Cooperative Information Systems Proc., 1999, 115-126
- [Veen92] Van Veen EA, Wortmann JC: Generative bill of material processing systems, Production Planning and Control, 1992, 3(3): 314-326
- [Voss06] Voss A: Klinische Behandlungspfade digital, Healthcare IT, 2006, <http://www.cio.de/healthcareit/bestpractice/827011/index.html>, abgerufen am 10.11.2006

Simulation von Grid-Versicherungen

Eine Analyse der Risikoallokation in Grid-Netzwerken

Autor: Werner Streitberger

Lehrstuhl für Wirtschaftsinformatik
Universität Bayreuth
Universitätsstrasse 30, 95440 Bayreuth
Werner.Streitberger@uni-bayreuth.de

Betreuer: Prof. Dr. Torsten Eymann

Lehrstuhl für Wirtschaftsinformatik
Universität Bayreuth
Universitätsstrasse 30, 95440 Bayreuth
Torsten.Eymann@uni-bayreuth.de

Abstract

Versicherungen bieten die Möglichkeit einer effizienten Risikoallokation in der realen Marktwirtschaft. Durch die Einführung künstlicher Märkte in Grid-Systemen zur markt-basierten Steuerung und Ressourcenallokation ist es möglich ökonomische Konzepte zur Risikoallokation anzuwenden. Dies ist notwendig, da durch die Anmietung von Diensten und Ressourcen an Fremdanbieter ein Risiko einhergeht, die vertraglich vereinbarten Garantien nicht oder nur unvollständig zu bekommen. Die Arbeit modelliert eine an das Grid-Netzwerk angepasste Versicherung zu effizienten Allokation von Risiken, die in Grid-Netzwerken auftreten. Mit Hilfe von Simulationen soll das Versicherungsmodell in einem Grid-Netzwerk evaluiert werden.

1 Einführung

1.1 Ausgangssituation und Problemstellung

Unternehmen müssen sich in einem Umfeld vielfältiger und neuer Anforderungen behaupten. Sich ständig wandelnde Herausforderungen, die in einer wachsenden Dynamik des Marktes

begründet sind, machen eine stetige Anpassung und Überprüfung des eingesetzten Wissens, der Technologie und insbesondere des eigenen Ressourceneinsatzes notwendig. In bestehenden Unternehmen hat sich der Einsatz von rechenintensiver IT (Informationstechnologie) bereits als für den Geschäftsbetrieb unverzichtbar erwiesen. Die Bereitstellung von aufwendiger IT gegenüber stehenden Kosten rechtfertigt jedoch nur selten die vollständige Abdeckung des potentiellen, maximal erwarteten Bedarfs von Software und Ressourcen, wie Speicher und Rechenleistung.

Eine Lösung ist das hinter dem Schlagwort „On-demand Computing“ stehende Konzept [FoKe99]. On-demand („bei Bedarf“) unterliegt der Vorstellung, nur einen Teil der IT bzw. IT-Dienste selbst bereit zu stellen und den variierenden, zusätzlichen Bedarf von einem Fremdanbieter zu beziehen. Die Dienste werden über kurzfristige, elektronische Verträge bei Bedarf (on-demand) angemietet und in die eigenen, betriebsinternen Prozesse integriert. Für Unternehmen bleiben die Kosten transparent, da sie nur die Ressourcen bezahlen müssen, die sie auch verbrauchen. Die Abrechnung erfolgt nach einem „pay-per-use“ Modell.

Eine Realisierung des On-demand Computing Konzept bietet das Grid-Computing. Es verfolgt die Idee ungenutzte Rechenkapazitäten über ein Netzwerk zusammenzuschließen und diese Kapazitäten flexibel nutzbar zu machen. Das besondere Potential des Grid-Computing liegt darin, verteilte Anwendungen zu unterstützen, um damit die Bearbeitung von beispielsweise sehr rechenintensiven Aufgaben überhaupt bzw. mit deutlich gesenkter Zeitdauer zu ermöglichen.

Eine Charakteristik des Grid-Computing ist die redundante, verteilte Bereitstellung und Zugriff auf Daten-, Rechendienste oder allgemein Anwendungsdienste, die sich über mehrere physikalisch getrennte Organisationen erstrecken können [FoKe99]. Dabei bleibt die Heterogenität der Dienste und Ressourcen dem Benutzer verborgen und er nimmt diese möglichst als homogen wahr. Ein Focus bei der Entwicklung zukünftiger Grid-Infrastrukturen liegt bei der effizienten Ausnutzung von verfügbaren Diensten und Ressourcen in dem Netzwerk. Aus diesem Grund ist es eine Schlüsselanforderung an Grid-Infrastrukturen, skalierbare und sich der Umgebung anpassende Allokationsmechanismen zu unterstützen. Jedoch sind aktuelle Grid-Systeme – häufig unter Verwendung einer Service-orientierten Architektur (SOA) – lose verbundene Systeme, die von Natur aus sehr dynamisch sind, was neue Herausforderungen an Allokationsmechanismen stellt [MoBa03].

Während aus technischer Sicht die Risiken der Akquise bereits mit der SOA angegangen und Lösungsansätze dargestellt wurden, ist auf der wirtschaftlichen Ebene noch ein Bedarf an Verfahren und Mechanismen notwendig, die den Erwerb von Diensten on-demand unterstützen und eine Reduktion des wirtschaftlichen Risikos bewirken. Hier sind zunächst die Dienstfindungs-, Dienstausswahl- und Wartezeiten bis zum Dienstzugang anzusiedeln, die für das Dienst suchende Unternehmen antizipierbar sein müssen, um per SLA (Service Level Agreement) verhandelte und zugesagte Garantien aufrechterhalten zu können.

Durch das dynamische Verhalten der Umgebung und der Delegation der Bereitstellung von Diensten und Ressourcen an einen Anbieter geht ein Risiko einher, die vertraglich verhandelten und vereinbarten Garantien unvollständig zu bekommen. Wenn die zugesicherten Dienste nicht geliefert werden, d. h. es kommt zu keiner Dienstaussführung auf den Grid Ressourcen, entsteht der größte Schaden. So sind die Garantien in den Verträgen erst während oder nach Dienstonutzung überprüfbar. Dabei stellt sich in der Dienstausswahl die Frage nach den Kriterien, nach denen der Konsument einen Anbieter selektiert; welches Risiko bzw. welche Risikokosten ist der Konsument bereit einzugehen und wie hoch sind die Kosten bei einer erneuten Allokation. Dabei soll mögliches Fehlverhalten (z. B. das Verletzen des vereinbarten Service Level Agreements) der Marktakteure in zukünftige Transaktionen mit einbezogen werden.

Es wird dabei angenommen, dass die Ressourcenallokation in einer markt-basierten Umgebung – häufig auch als Grid-Markt oder Grid-Ökonomie bezeichnet – durchgeführt wird, die es Konsumenten und Anbietern erlaubt, Dienste und Ressourcen zu kaufen, zu verkaufen und deren SLA zu verhandeln, wobei die Anwesenheit eines Marktes die Verwendung von ökonomischen Konzepten erlaubt [ENRS06]. Markt-basierte Kontrolle wird als Ressourcenallokations- und Steuerungsverfahren in Analogie zu einer freien Marktwirtschaft eingesetzt. Dies erfolgt mit dem Ziel, die gleiche Dezentralität, Robustheit und Selbstorganisation zu erreichen, die real existierende Ökonomien zeigen. In der künstlichen Grid-Ökonomie existiert bisher keine effiziente Allokation der oben beschriebenen Risiken.

In real existierenden Ökonomien sind verschiedene Verfahren bekannt, mit denen Risiken bewältigt werden können. Diese Aufgabe übernehmen häufig Versicherungen als Instrument ein bestimmtes Risiko zu begrenzen. Die Versicherung ist eine vertrauenswürdige Institution, die das Risiko einer Vertragsverletzung begrenzt. Versicherungen folgen dem Grundprinzip der kollektiven Risikoübernahme: Viele zahlen einen Versicherungsbetrag in den Geldtopf der Versicherung ein, um beim Eintreten des Versicherungsfalls aus diesem Geldtopf einen

Schadenausgleich zu erhalten. Da der Versicherungsfall nur bei wenigen Versicherten eintreten wird, reicht der Geldtopf bei bezahlbarem Beitrag aus. Wenn Versicherungen viele Versicherungsnehmer haben, dann können sie die Risiken breit verteilen und damit das Risiko des Einzelnen deutlich zu verringern. Sind die Risiken unabhängig von einander und die Anzahl der Versicherungsnehmer groß, so können die individuellen Risiken effizient durch ein Versicherungsunternehmen beseitigt werden [Milg92]. Der Inhalt des Versicherungsvertrags ist ein Schutzversprechen des Versicherers, nach einem konkreten Schadensfall eine vertraglich vereinbarte Leistung zu erbringen. Diese Leistung wird durch monetäre Kompensation erbracht. Voraussetzung für den Einsatz einer Versicherung ist, dass der Umfang der Schäden statistisch abschätzbar ist und demnach mit versicherungsmathematischen Methoden der von jedem Mitglied des Kollektivs benötigte Beitrag bestimmbar ist. Mögliche Schäden im Grid-Computing sind die Nichteinhaltung einer vereinbarten Dienstgarantie, Abbruch der Dienstaufführung oder der Ausfall von Teil-Ressourcen. Diese lassen sich durch eine Zufallsvariable, die das Risiko repräsentiert, beschreiben. Jedoch gibt es auch extreme Ausprägungen wie den Ausfall aller zur Verfügung stehender Ressourcen. Diese gelten als nicht versicherbar und werden in der Arbeit nicht weiter betrachtet.

Versicherbare Risiken sollen durch die Einführung einer Grid-Versicherung effizient abgewickelt werden. Der Aufwand zur Allokation des Risikos soll dabei möglichst gering sein. Dies bedeutet, dass die zugrunde liegende Infrastruktur des Grid Systems möglichst nur in geringen Maß belastet werden soll. Die allgemeinen Funktionen der Versicherung als Institution in einer Marktwirtschaft werden im nächsten Kapitel dargestellt und daraus die Ziele der Arbeit abgeleitet.

1.2 Zielsetzung und Überblick

Allgemein fördert in einer Ökonomie eine Versicherung die wirtschaftliche Effizienz auf mindestens sechs verschiedenen Wegen [ZwEi03]:

- **Verbesserung der Risikoallokation:** Die effiziente Allokation von Risiken verringert die Transaktionskosten und reduziert die Schäden durch prompte Schadensregulierung und technische Kontrollen. Beides steigert die Effizienz des Wirtschaftsprozesses und trägt zum Wachstum und Stabilisierung der Wirtschaft bei.

Übertragung auf Grid-Märkte:

Transaktionskosten in Grid-Märkten fallen in der Dienstfindung, der Dienstauswahl, bei Wartezeiten bis zum Dienstzugang und bei der Bewertung und Sanktionierung der aktuellen Transaktion an. Sie legen die zusätzlichen Kosten der Dienstnutzung fest und können als Zeit- oder Infrastrukturkosten gemessen werden. Als Empfehlungsdienst kann eine Versicherung das zukünftige Transaktionsverhalten bewerten und damit die Transaktionskosten bei einer Fehlallokation zu verringern. Es ist ebenfalls möglich, ein Fehlverhalten des Anbieters nach Dienstausführung durch Erhöhung der Versicherungsprämie sanktionieren. Dazu muss der Anbieter durch die Versicherung überwacht werden, um eine mögliche Vertragsverletzung feststellen zu können. Auch eine prompte Schadenregulierung durch Transfer des versicherten Schadens an den Versicherungsnehmer kann in Grid-Märkten umgesetzt werden.

- **Vermögenssicherung:** Der Schutz bestehenden Vermögens ist eine bedeutende Funktion der Versicherung deshalb, weil damit den Wirtschaftssubjekten eine verbesserte Grundlage für ihre Planung bereitgestellt wird. Oft ermöglicht erst diese gesicherte Grundlage, risikobehaftete, aber gewinnträchtige Geschäfte in Angriff zu nehmen.

Übertragung auf Grid-Märkte:

Ein Konsument auf dem Grid-Markt hat durch die Einführung einer Versicherung die Möglichkeit den maximalen Schaden seiner Transaktion zu versichern. Dies ermöglicht ihm risikobehaftete Ressourcenanbieter auszuwählen und bei Vertragsverletzung seinen entstandenen Schaden erstattet zu bekommen. Dabei wird angenommen, dass der Konsument risikoavers handelt und zur Vermögenssicherung für jede Allokation eine Versicherung abschließt. Er wählt bei gleichen Angeboten den Anbieter aus, der das geringste Risiko bezüglich des Ergebnisses ausweist. Schlägt die Transaktion fehl, so ist dem Konsumenten kein finanzieller Schaden entstanden und er hat die Möglichkeit eine erneute Transaktion zu starten.

- **Kapitelakkumulation:** Versicherung hat eine Kapitalakkumulationsfunktion. Durch den Kauf von Versicherungsschutz entsteht notwendig ein Kapitelfond, weil die Prämien von Versicherungen zu Beginn der Versicherungsperiode bezahlt

werden und so einen Fonds bilden, aus dem im Laufe des Jahres die Entschädigungen und die Aufwendungen für den Versicherungsbetrieb bezahlt werden. So ergibt sich ein ständig vorhandener Fond, der im Durchschnitt größer ist als die Hälfte der jährlichen Prämieinnahmen.

Übertragung auf Grid-Märkte:

In der Grid-Ökonomie berechnen die Versicherungen ihre Prämien für eine festgelegte Periode. Eine Fondsbildung über ein Jahr, wie es häufig in real existierenden Versicherungen verwendet wird, ist nicht realistisch. Der Zeitraum der Periode muss kürzer gewählt werden, um auf Änderungen in der Grid-Ökonomie schneller mit angepassten Prämien reagieren zu können.

- **Mobilisierung von finanziellen Ressourcen:** Diese beruht einerseits darauf, dass über den Kauf von Versicherungsschutz die Spartätigkeit angeregt wird. Individuell betrachtet jeder Versicherte seine Prämie als Sparbeitrag, der sich ansammelt und später fällig wird. Gesamtwirtschaftlich betrachtet trägt die Versicherung mit ihrer Kapitalansammlung in ihrer Aufbauphase zur Kapitalbildung bei.

Übertragung auf Grid-Märkte:

Auch die Mobilisierung von finanziellen Ressourcen, wie sie häufig bei Lebensversicherungen verwendet wird, ist auf Grid-Märkten nicht anwendbar, da keine finanziellen Anlagemöglichkeiten für die Versicherungsbeiträge existieren. Die Umsetzung einer Grid-Versicherung konzentriert sich aus Gründen der Vereinfachung auf den Risikotransfer. Das Anlagegeschäft, das eine real existierende Versicherung zusätzlich durchführt, wird nicht weiter betrachtet.

- **Kontrolle des Unternehmensverhaltens:** Versicherung dient auch der Kontrolle des Unternehmensverhaltens. Denn üblicherweise werden Versicherungen risikogerecht kalkuliert, d.h. die Prämie ist umso höher, je größer das Risiko (Schadensbetrag und/oder Schadenswahrscheinlichkeit). Dieses Verursacherprinzip schafft Anreize für die (produzierenden) Unternehmen, die von Ihnen ausgehenden Risiken zu reduzieren, insoweit Risikoreduktion billiger ist als

die Versicherung des Risikos. Folglich üben Versicherer eine gewisse Kontrolle auf das Verhalten aus, die sicherstellt, dass die Ressourcen den produktivsten Verwendungszweck zugeführt werden und nicht verschwendet werden.

Übertragung auf Grid-Märkte:

Durch die Kalkulation von risikogerechten Prämien werden Transaktionen mit unzuverlässigen Dienst- oder Ressourcenanbieter auf Grid-Märkten für den Konsumenten durch die hohen Versicherungsprämien sehr teuer. Dadurch werden diese Anbieter bei der Auswahl als zukünftige Transaktionspartner unter Annahme von Risikoaversion weniger häufig bevorzugt.

- **Entlastung des Staates:** Auch die Entlastung des Staates ist eine wichtige Funktion der Versicherung. Durch den Kauf einer Versicherung wird das Gemeinwesen insofern entlastet, als es ohne Versicherung für Schäden einzustehen hätte, die einzelne Mitglieder der Gesellschaft erleiden und die diese nicht selbst tragen können. Dies bezieht sich auf den Einzelnen, der einen Schaden erleidet, aber auch Dritte, die Opfer sein können. Für diese Schäden müsste der Staat aufkommen.

Übertragung auf Grid-Märkte:

Die Funktion des Staates findet sich auf Grid-Märkten in der gemeinsam benutzten technischen Infrastruktur. Durch eine Reduzierung der Transaktionskosten, ist es möglich die Infrastruktur weniger zu belasten. So müssen weniger Nachrichten über das Kommunikationsnetzwerk verschickt werden, da durch die Empfehlungen der Versicherung mögliche Fehlallokationen vermieden werden. Eine weniger belastete Infrastruktur führt zu weniger Schäden an Dritten, welche die gemeinsame Infrastruktur verwenden.

Die ausgewählten Funktionen von Versicherungen in einer realen Marktwirtschaft werden in der Arbeit in einer künstlichen Grid-Ökonomie umgesetzt. Dabei wird untersucht, ob die beschriebenen Verhaltensmuster auf künstlichen Grid-Märkten auftreten und Versicherungen eine geeignete Möglichkeit sind, Risiken, die durch die Delegation einer Dienstaufführung an einen Fremdanbieter auftreten, zu verringern. Mit Hilfe von Simulationen in einem Grid-

Simulator wird das implementierte Versicherungsmodell analysiert. Diese stehen im Mittelpunkt der Arbeit.

Die Schäden treten unabhängig voneinander mit einer bestimmten Wahrscheinlichkeit auf und bilden eine Schadenverteilung. Die Häufigkeit und Höhe der Schäden wird in den Simulationen verändert. Einzelne Ressourcenallokationen werden versichert und das Verhalten der Versicherung während der Simulation durch Metriken gemessen. Simulationen bieten eine schnelle Evaluierung des Versicherungsmodells in verschiedenartigen Grid-Netzwerken. Zusätzlich erlaubt der Einsatz von Simulationen, die Umgebungsparameter wohl definiert zu ändern.

Im nächsten Kapitel wird das zu simulierende Modell näher betrachtet. Das Grid-Szenario wird vorgestellt, das Versicherungsmodell beschrieben und der Aufbau der Simulationen erläutert.

2 Eine transaktionsbasierte Versicherung für Grid-Märkte

Dieses Kapitel stellt den aktuellen Stand der Arbeit vor. Kapitel 2.1 beschreibt detailliert das Modell des Grid-Markts und leitet daraus die Anforderungen an eine Grid-Versicherung ab. Das Design einer Grid-Versicherung schildert Kapitel 2.2. Kapitel 2.3 schließt mit einer kurzen Einführung über den gewählten Grid-Simulator ab, mit dessen Hilfe Simulationen des Grid-Versicherungsmodells durchgeführt werden.

2.1 Analyse der Umgebung und Anforderungen an eine Grid-Versicherung

Wie in der Einleitung bereits erwähnt bieten Serviceorientierte Architekturen eine Möglichkeit für die Umsetzung einer Grid-Computing Infrastruktur, auf der Teilnehmer Anwendungsdienste anbieten und nachfragen. Aus diesem Grund haben sich Grid-Systeme zunehmend als Netzwerke miteinander agierender Dienste herausgebildet (Siehe Abbildung 1).

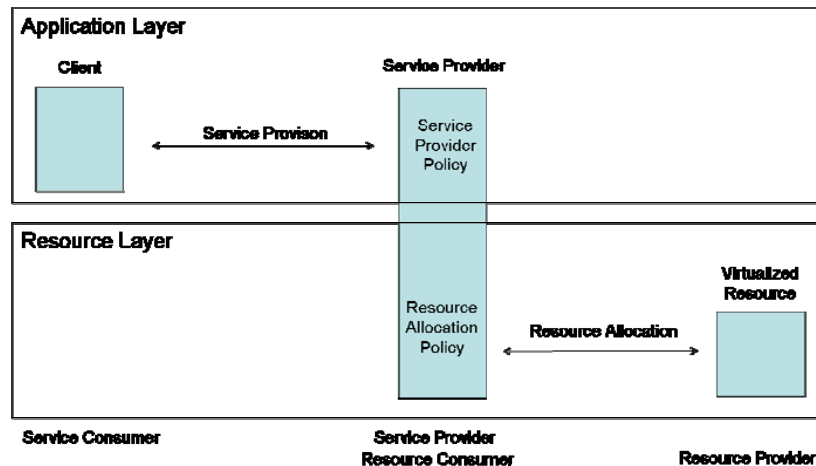


Abbildung 1: Modell der Grid-Umgebung

Dies erlaubt Serviceprovider sich auf die Bereitstellung von Diensten und die Allokation der benötigten Ressourcen zu konzentrieren, während sich die Organisationen auf ihre zentralen Geschäftsaktivitäten fokussieren. In diesem Modell nutzen die Clients Dienste ohne Wissen über die interne Infrastruktur auf der Ressourcenebene. Sie verhandeln über die minimal benötigte Servicequalität und möchten Garantien über die Bereitstellung von Diensten. Diese Garantien werden verhandelt und zwischen Dienstanbieter und Dienstkonsument durch Service Level Agreements (SLAs) festgelegt. Die Policy eines Dienstbieters bestimmt, wie die Bereitstellung der Service durchgeführt wird, welche Ressourcen für jeden Dienst benötigt werden und ob neue Ressourcen reserviert werden müssen, um Bedarfsspitzen zu decken.

Die Einführung von virtuellen Maschinen und einer Ressourcenvirtualisierung erlaubt die Verwendung von verschiedenen Policies für die Bereitstellung von Diensten und die Allokation von Ressourcen. Damit ein Serviceanbieter die SLAs dem Client gegenüber einhalten kann, muss er Schäden, die durch nicht bereitgestellte oder nur teilweise verfügbare Ressourcen auf dem Ressourcenmarkt, möglichst gering halten. Er benötigt eine effiziente Allokation von Risiken auf dem Ressourcenmarkt, um seine Transaktionskosten reduzieren zu können. Unter einer Transaktion aus dem Ressourcenmarkt versteht man die Allokation eines Ressourcenbündels. Ein Ressourcenbündel besteht aus Einzelressourcen, die zusammen nachgefragt werden. Ist eine Ressource des Ressourcenbündels nicht verfügbar, tritt ein Schaden ein. Dieser Schaden kann durch eine Versicherung versichert werden. Wichtig für die Berechnung einer Versicherungsprämie ist die Unabhängigkeit der Schadensereignisse. Dies kann garantiert werden, wenn die Ressourcen exklusiv einem Nachfrager zur Verfügung gestellt

werden, wie es häufig im Bereich High-Performance Computing und On-Demand Computing eingesetzt wird. Ein Ressourcenbündel auf dem Ressourcenmarkt sieht wie folgt aus: (Anzahl Prozessoren 4, Hauptspeicher 16GB, Festplattenplatz 300GB). Für die Ausführung eines Dienstes werden 4 Prozessoren, 16 GB Hauptspeicher und 300GB Festplattenplatz benötigt. Diese Anforderungen muss ein Ressourcenanbieter erfüllen, um den Nachfrager die benötigten Ressourcen für die Dienstauführung bereitzustellen. Die einzelnen Teilressourcen besitzen dabei eine Ausfallwahrscheinlichkeit, die zu einem Versicherungsschaden führen. Durch das Verursacherprinzip muss ein Ressourcenanbieter damit rechnen, ein höheres Risiko aus Sicht der Versicherung darzustellen, wenn er unzuverlässige Ressourcen mit einer hohen Ausfallwahrscheinlichkeit bereitstellt. Dies schafft Anreize für Ressourcenanbieter, die von ihm ausgehenden Risiken zu reduzieren, damit die Risikoabsicherung für einen Ressourcenanbieter günstig ist und er den Anbieter bei gleicher Leistung bevorzugt.

Abbildung 2 zeigt schematisch die Aufgaben der Versicherung in dem beschriebenen Grid-Ressourcenmarkt. Ressourcenkonsument und Ressourcenanbieter haben einen Vertrag über die Nutzung der Ressourcen abgeschlossen. Die Versicherung kontrolliert den Anbieter während der Ausführung des Dienstes und schützt das bestehende Vermögen des Konsumenten im Schadensfall. Aus Vereinfachungsgründen ist die Versicherung ein Monopolist auf dem Markt. Es gibt keine konkurrierenden Versicherungen mit dem gleichen Angebot.

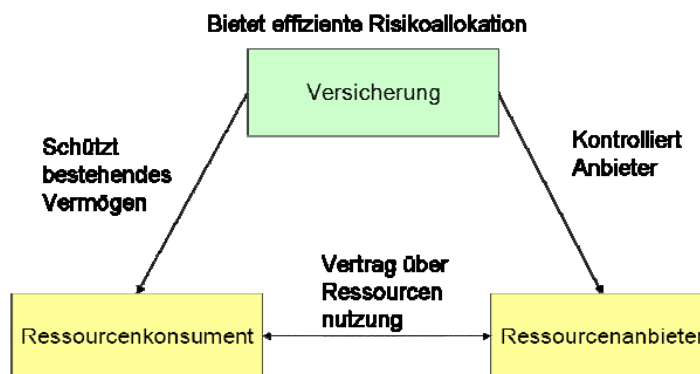


Abbildung 2: Schematischer Aufbau des Ressourcenmarkts

Im Detail zeigt Abbildung 3 die Aufgaben der Versicherung während einer Transaktion. Die Versicherung kann Aufgaben der technischen Kontrolle des Anbieters übernehmen. Dies ist jedoch nicht zwingend erforderlich. Die Leistungen können auch von einer anderen Instanz erbracht werden. Wichtig für die Versicherung ist die Schadensverteilung, die Grundlage einer

risikogerechten Prämienkalkulation ist. In dieser werden die erfolgreichen Transaktionen, sowie Schäden protokolliert und aggregiert. Da die Versicherung auf dem Ressourcenmarkt nicht nur ein Produkt versichert, sondern verschiedene Ressourcenbündelkombinationen, wird für jedes dieser Güterbündel eine Schadenstatistik verwaltet. Häufig wird auch von einem Risikoportfolio gesprochen, das die Versicherung anbietet.

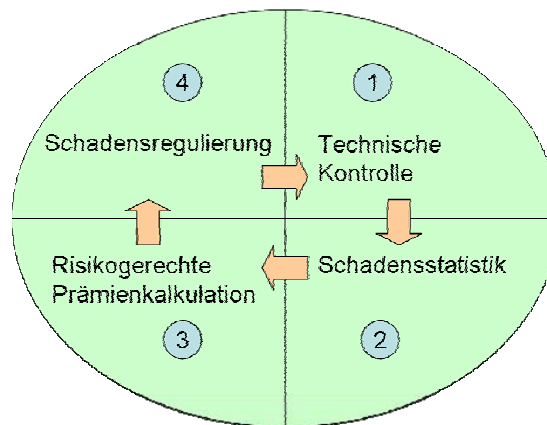


Abbildung 3: Aufgaben einer Versicherung für Grid Systeme

Die risikogerechte Prämienkalkulation berechnet den Preis zu dem das Versicherungsprodukt angeboten werden kann. Das Ziel ist es eine Preisuntergrenze zu bestimmen, bei der es sich gerade noch lohnt, ein einzelnes Versicherungsprodukt anzubieten. Diese Preisuntergrenze hängt in der traditionellen Prämienkalkulation von den Eigenschaften der Schadenverteilung sowie einer exogen vorgegebenen Wahrscheinlichkeit der Solvenz ab. Hier kommen Elemente der Wahrscheinlichkeitstheorie zur Anwendung, die im folgenden Kapitel näher erläutert werden.

Als letzte Aufgabe der Versicherung lässt sich die Schadenregulierung identifizieren. Hier werden die abgesicherten Vermögen an die Versicherungsnehmer, den Ressourcennachfrager ausbezahlt. Das gewählte Design der Versicherung und eine mögliche Implementierung wird im nächsten Abschnitt vorgestellt.

2.2 Design und Implementierung der Versicherung

Die Risikotheorie der Versicherung konzentriert sich vor allem auf den Risikotransfer des Versicherungsunternehmens. Im Zentrum der Grid-Versicherung steht die Beschreibung und Voraussage der Verpflichtungen, die sich aus der Risikozeichnung ergeben. Bei Eintreten des Versicherungsfalles entsteht gegenüber dem Versicherungsunternehmen eine

Zahlungsforderung und es muss den vertraglich vereinbarten Betrag an den Versicherungsnehmer auszahlen. Jedoch kommt es nicht bei jedem Vertrag des Unternehmens zwangsläufig zur Auszahlung und die Höhe der Auszahlung hängt von dem tatsächlich eingetretenen Schaden ab. Der verursachte Schaden jedes Vertrages kann als ein Wert betrachtet werden, der dem Zufall unterworfen ist. Allgemein treten die Schäden im Zeitablauf unregelmäßig und in unterschiedlicher Höhe ein. Sie bilden einen stochastischen Prozess, der aus zwei Komponenten besteht [ZwEi03]:

- Unsichere Zahl der Schäden
- Unsichere Höhe des jeweiligen Schadens.

Die unsichere Zahl der Schäden lässt sich auf einem Grid-Markt auf die Nichterfüllung der vereinbarten Garantien zurückführen. Diese können beispielsweise durch den Ausfall einer Ressource hervorgerufen werden. Die unsichere Höhe des jeweiligen Schaden beruht auf dem Marktpreis zu dem das Ressourcenbündel gehandelt wird. Knappe Ressourcen führen zu einem höheren Marktpreis zu dem das Gut gehandelt wird. Dies erhöht auch die Schadenssumme, die die Versicherung bei einer Fehlallokation ausbezahlen muss.

Die Prämienkalkulation für Einzelrisiken erfolgt mit Hilfe eines Funktionals p , das der Zufallsvariablen X eine reelle Zahl zuordnet. Dabei hängt das Prämienkalkulationsprinzip $p(X)$ von der Schadenverteilung ab. Wichtig für die Berechnung der Prämien sind der Erwartungswert und die Varianz der Schadenverteilung. Ausgangspunkt für die praxisnahe Berechnung der Prämien ist das Äquivalenzprinzip. Die Risikoprämie entspricht dabei dem Erwartungswert der Schadenverteilung. Dies würde jedoch einen schnellen Ruin der Versicherung bedeuten, da sie keinen Risikozuschlag kalkuliert hat. Das Erwartungswertprinzip erweitert das Äquivalenzprinzip um einen beliebig wählbaren Risikozuschlag, der verhindern soll, dass die Versicherung insolvent wird. Der Risikozuschlag wird so gewählt, dass die Versicherung mit einer bestimmten Wahrscheinlichkeit keinen Ruin erleidet. Ebenfalls auf dem Erwartungswert bauen das Varianz und Standardabweichungsprinzip auf, die als alternative Prämienkalkulationsprinzipien für Grid-Versicherungen evaluiert werden. Um der Versicherung die Möglichkeit zu geben den Vorteil einer verbesserten Risikosituation an den Versicherungsnehmer weiterzugeben, wird der Risikozuschlag angepasst: Je höher die Sicherheit, desto niedriger, die vom Versicherungsunternehmen geforderte Prämie.

Die Bestimmung der Sicherheit bei der Bereitstellung lehnt sich das Modell nutzungsabhängiger Autoversicherungstarife an, das seit Oktober 2006 die britische Norwich Union ihren Kunden anbietet [Thol06]. Das Prämienmodell verfolgt dem „pay as you drive“ Grundsatz. Dies bedeutet, dass die zu bezahlende Versicherungsprämie abhängig von der Autonutzung ist. Die Prämie wird dabei auf Basis von drei Parametern berechnet: Tageszeit, gefahrener Straßentyp und Kilometerzahl. Untersuchungen haben ergeben, dass sich die Unfallrate in diesem Tarif deutlich gesunken ist und die Autofahrer durch die gesunken Schäden weniger Versicherungsprämie bezahlen mussten. Der Tarif zahlte sich vor allem für Wenigfahrer aus.

Im Gegensatz zu dem Autoversicherungsmodell, hängt die Versicherungsprämie in der Grid-Versicherung nicht von dem Verhalten des Ressourcenkonsumenten ab, sondern von der Leistung des Ressourcenanbieters. Überträgt man dieses Modell auf Grid-Versicherungen, so lassen sich durch eine nutzungsabhängige Prämienberechnung die Risikokosten für den Konsumenten der Ressourcen genauer kalkulieren. Als Parameter, die die Nutzung der Grid Ressourcen beschreiben, bieten sich die Häufigkeit der getätigten Transaktionen an, die Nutzungsdauer der Ressourcen und die Auslastung des Systems.

Um diese Prinzipien auf eine Grid-Versicherung anzuwenden, muss eine Periode festgelegt werden, in der die Schadenverteilung bestimmt wird. Daraus kann der Erwartungswert und die Varianz abhängig von der Schadenverteilung für ein Einzelrisiko berechnet werden. Diese bilden neben den nutzungsabhängigen Komponenten die Grundlage der Prämienberechnung nach einem der oben genannten Prämienkalkulationsprinzipien für die nächste Periode.

2.3 Test des Modells durch Simulation

Die Grid-Versicherung wird in einem Grid Simulator implementiert. Dieser simuliert das in Abbildung 1 dargestellte Szenario. Die einzelnen Akteure sind als Software-Agenten implementiert. Diesem Szenario wird die Grid-Versicherung als weiterer Software Agent hinzugefügt, der automatisiert das oben beschriebene Versicherungsmodell realisiert. Die Akteure kommunizieren durch den Austausch von Nachrichten miteinander. Jedem Ressourcenanbieter wird eine Ausfallwahrscheinlichkeit zugewiesen, die bei Start der Simulation festgelegt werden und sich während der Simulation nicht ändern.

Das Grid-Versicherungsmodell wird dabei mit drei Parametern initialisiert:

- **Überschuss zu Beginn der Simulation:** Am Anfang der Simulation benötigt die Versicherung ein Budget, mit dem es die bei Beginn auftretenden Schäden auszahlen kann. Der Prämieeneinnahmeprozess der Versicherung wird den Überschuss während der Simulation festlegen. Dieser Überschuss wird als Metrik gemessen und am Ende der Simulation ausgewertet.
- **Höhe des Risikozuschlags:** Dieser steuert den Prämieeneinnahmeprozess und muss dazu so gewählt werden, dass die Versicherung nicht in den Ruin getrieben wird. Dieser Zuschlag ändert sich während der Simulation nicht. Anpassung der Prämie erfolgt durch die oben beschriebene nutzungsabhängige Komponente der Versicherung.
- **Periode der Schadenverteilung:** Dieser Wert legt die Bewertungsgrundlage für die Versicherung fest. Wird er zu groß gewählt, ist es der Versicherung nicht möglich schnell auf die sich ändernde Risikolage zu reagieren durch die Berechnung von neuen Prämien auf Basis des Erwartungswerts und der Varianz der Schadenverteilung. Eine zu kleine Periode würde eine häufige Neuberechnung der Schadenverteilung und der Prämien führen, was einen erhöhten Rechenaufwand durch die Versicherung darstellt.

Die Umsetzung des Versicherungsmodells in den ausgewählten Grid Simulator erfolgt durch die Umsetzung der in Abbildung 3 gezeigten vier Schritten. Ziel der Simulationen ist es Muster zu finden, die den in Kapitel 2 aufgestellten Funktionen real existierender Versicherungen entsprechen.

3 Zusammenfassung und Ausblick

Die Arbeit überträgt das Modell der Versicherungen als Möglichkeit der effizienten Risikoallokation. Bei der Umsetzung in ein Grid-Netzwerk, werden Grid-spezifische Eigenschaften berücksichtigt. In Anlehnung an das „pay-as-you-drive“ Modell neuerer Autoversicherungen, wird das aus dem On-Demand Computing bekannte Modell „pay-per-use“ um einen Risikozuschlag erweitert, der dem Ressourcenkonsument die Möglichkeit gibt das Risiko einer Allokation anzuschließen. Die Grid-Versicherung verwendet dabei einfache und

praxisnahe Konzepte aus der Risikotheorie zur Kalkulation der Prämien, um eine schnelle und automatisierte Berechnung zu ermöglichen.

Die nächsten Schritte in der Arbeit sind die Verfeinerung des Modells und die anschließende Implementierung des Grid-Versicherungsmodell als Software-Agent im vorgestellten Szenario. Weitere Arbeiten zur Analyse der Schadenverteilungen der Grid Ressourcen sind notwendig, um ein Prämienkalkulationsprinzip auswählen zu können.

Literaturverzeichnis

- [EACC05] Eymann, T.; Ardaiz, O.; Catalano, M.; Chacin, P.; Chao, I.; Freitag, F.; Gallegati, M.; Giulioni, G.; Joita, L.; Navarro, L.; Neumann, D.; Rana, O. F.; Reinicke, M.; Schiaffino, R.; Schnizler, B.; Streitberger, W.; Veit, D.; Zini F.: Catallaxy-based grid markets. In Proceedings of the First International Workshop on Smart Grid Technologies (SGT05), 2005.
- [ENRS06] Eymann, T.; Neumann, D.; Reinicke, M.; Schnizler, B.; Streitberger, W.; Veit D.: On the Design of a Two-Tiered Grid Market Structure. Business Applications of P2P and Grid Computing, MKWI, 2006.
- [FoKe99] Foster, I.; Kesselman, C.: The Grid: Blueprint for a New Computing Infrastructure. Morgan Kauffman, San Francisco, 1999. 2, 29, 223
- [Fran03] Frank, M.: "On-Demand-Computing ist eine Managementaufgabe". Forschungsbericht, Computer Zeitung, Abruf am 11. November 2003, 2003. 2
- [Milg92] Milgrom, P.: „Economics, organizations, and management“, Prentice-Hall, 1992
- [MoBa03] Montresor, A.; Baboglu, O.: Biology-inspired approaches to peer-to-peer computing in bison. In Proceedings of the 3rd International Conference on Intelligent System Design and Applications (ISDA'03), Advances in Soft Computing, pages 515–522. Springer-Verlag, Aug. 2003.

- [RSEC06] Reinicke, M.; Streitberger, W.; Eymann, T.; Catalano, M.; Giulioni, G.: Economic Evaluation Framework of Resource Allocation Methods in Service-Oriented Architectures Proceedings of the 8th Conference on E-Commerce Technology (CEC06), San Francisco, June 26th-29th
- [Shan04] Shankland, S.: "IBM: On-demand computing has arrived". Forschungsbericht, CNET News.com, URL [http://news.com.com/IBM: +On-demand+computing+has+arrived/2100-7784_3-5106577.html?part=business2-cnet](http://news.com.com/IBM:+On-demand+computing+has+arrived/2100-7784_3-5106577.html?part=business2-cnet), 2004. 2, 184
- [Thol06] Thole, S.: "KFZ-Versicherung passt sich dem Gebrauch an". Computerzeitung Nr. 47, 20. November 2006
- [ZwEi03] Zweifel, P.; Eisen, R.: „Versicherungsökonomie“, 2., verb. Auflage, Springer-Verlag, 2003.

Incentive Engineering and Transaction Pricing in the Securities Trading Value Chain

Matthias Burghardt

Institute of Information Systems and Management (IISM)
Information & Market Engineering
Universität Karlsruhe (TH)
76128 Karlsruhe
matthias.burghardt@iism.uni-karlsruhe.de

Supervisor:

Prof. Dr. Christof Weinhardt

Institute of Information Systems and Management (IISM)
Information & Market Engineering
Universität Karlsruhe (TH)
76128 Karlsruhe

Abstract

Most of the research on transaction costs in the market microstructure literature focuses on implicit transaction costs. Research on the design of price schedules for explicit transaction fees is rare. This paper analyzes and classifies different price schedules and discusses their application to the market transaction business. The discussion highlights design issues and the need for a structured approach for price schedule design in the context of market engineering.

In order to get some insights into customer order behavior, we conduct a trading experiment where participants trade virtual stocks on an electronic platform within a time period of three weeks. During three weeks, participants face transaction fees of different types. Order frequency and volume is measured and related to the price schedule in place. We find that both variables are influenced by transaction fees. We also try to identify price elasticities for groups with different income and use this information for a structured approach towards a nonlinear price schedule design.

1 Introduction

1.1 Motivation

In recent years, the global exchange landscape has been undergoing a constant change: Trading venues where trades have taken place in open outcry auctions are replaced by electronic trading systems. Continuous trading usually replaces trading via intermediaries like specialists who determine prices and supply liquidity in single securities. Some traders fear that market quality could decrease if human judgment by intermediaries in the securities trading value chain is replaced by automatic matching algorithms. In some markets, e.g. the New York Stock Exchange (NYSE), or in some types of securities, e.g. covered warrants, stock exchanges still use some kind of specialist system in order to guarantee a certain order execution quality. In coming years, market operators will have to decide whether there is still need for human intermediaries.

Another development which has influenced the electronic value chain in securities trading in recent years is the use of order routing systems and market access infrastructure which is provided to retail customers by online brokers. For the first time, retail investors get direct electronic access to trading venues world wide. The increasing trading volume and the decreasing costs for transaction services at the same time require ongoing disintermediation of processes. Brokerages and market operators must develop innovative business and pricing models to stay competitive in the long run.

1.2 Objective and Outline

This paper is organized as follows: In section 2, an introduction to transaction fee models is presented. A motivation, related research and a formal representation of transaction fees are discussed. Furthermore, some real-life examples from stock exchanges are given. In section 3, an experiment on transaction fees is introduced and some first results are outlined. Section 4 concludes with an outlook on further research questions.

2 Transaction Fee Models

2.1 Motivation

Pricing of transaction services in electronic financial markets has been undergoing a continuous change. Recently, the New York Stock Exchange announced a major change in their transaction fee schedule which took effect on August 1, 2006. This initiative represented the first significant price change in many years, most probably driven by the acquisition of Archipelago, the development of the NYSE Hybrid market, and NYSE's attempts to foster electronic trading.

When the traditional New York Stock Exchange announced its new price schedule, so too did the fully-electronic NASDAQ Stock Market and changed the pricing scheme for NYSE-listed securities. Transaction fees seem to play an important strategic role for exchanges and marketplaces.

Recent developments in the design of price schedules such as "free trade" and "no fee" offers by banks and brokerages emphasize the relevance of research on the structured engineering of price schedules and customer price sensitivity regarding transaction services. Obviously, customers are aware of different price schedules and take them into consideration when they decide about placing their order.

In general, price schedules of transaction services may depend on several variables, including the value and the volume of the order, both on a per-trade and a per-period basis. Another variable that may be employed is the sum of transaction fees already paid in a specific period of time. Depending on these variables, price schedules can be classified into different classes of schedules, differing in complexity, depth, and therefore ability for price discrimination. The classification and discussion of price schedules and is regarded as a contribution towards the structured approach of price schedule design in the context of market engineering [WeHN03].

2.2 Related Research

Much work has been done around transaction costs in financial markets but most of the literature focuses on implicit transaction costs like the spread or the market impact. Commissions and fees are usually neglected.

Wilson [Wils93] published a monograph about nonlinear pricing and tariff design in the context of electricity and telephone markets which has become a standard reference in the economic pricing literature. Mathematically founded, he showed how tariffs should be designed and based his research on economic theory. For practitioners, this approach does not seem to be easily

applicable since the results are based on a set of assumptions which have to be estimated in order to find the optimal solution. Furthermore, Wilson does not take marketing and communication aspects into account.

Dolan [Dola87] integrates the economics and marketing literature and presents an overview of possible motivations for quantity discounts. He narrows his tariff design to three basic types which he compares with each other. The same has been done by Leland and Meyer [LeMe76] who compare two-part pricing and block pricing. Although Dolan presents some price schedule design issues for pricing managers, these could be further expanded and detailed to be an integral part of the market engineering approach.

In the German pricing literature, Hermann Simon [Simo82] presents the basic principles of price management in a broader marketing context. Tacke [Tack89] focuses on non-linear pricing especially with two-part and two-block schedules described in [Dola87]. He presents some good examples for the application of price discrimination and shows when to use which kind of price schedules. Skiera [Skie99] finally narrows the application domain to services and discusses price discrimination in that context.

2.3 Formal Representation of Transaction Fee Schedules

Suppliers of transaction services have several possibilities to charge for their services. The types of price schedules that are most commonly used are uniform schedules where the price is proportional to the quantity purchased, and nonuniform schedules where there are price breaks depending on the volume.

2.3.1 Types of Nonlinear Price Schedules

A price schedule specifies the relation between the marginal price per unit and the number of units per transaction. Most generally, a price schedule can be represented by a function $R(q)$ where $R(q)$ is the transaction fee for the order size q .

Nonlinear price schedules can be classified into two-part schedules, two-block schedules (which can be easily extended to n -block schedules), and all-units quantity discount schedules.

(i) Two-part price schedules consist of a fixed fee F and a constant marginal price p . Essentially, the customer pays the fixed fee F for the right to place an order at all, and a variable fee pq depending on the size q of the order. The two-part schedule can be written as:

$$R(q) = \begin{cases} F + pq, & q > 0 \\ 0, & q = 0 \end{cases}$$

(ii) Two-block price schedules consist of two different marginal prices p_1 and p_2 where p_1q is charged for an order size of up to x units. If the order is greater than x , the first x units are priced at p_1x , and all subsequent units are priced at p_2 per unit. This can be written as:

$$R(q) = \begin{cases} p_1q, & 0 \leq q \leq x \\ p_1x + p_2(q - x), & q > x \end{cases}$$

(iii) All-units quantity discount schedules consist of several different marginal prices which are applied for each unit depending on the total size of the order. That means if a certain quantity level is exceeded, the corresponding marginal price applies to all units. The mathematical formula for the all-units quantity discount schedule is as follows:

$$R(q) = \begin{cases} p_1q, & 0 \leq q < x \\ p_2q, & q \geq x \end{cases}$$

As a result of the different marginal prices associated with different quantities, all-units quantity discount schedules require that there are discontinuities in the $R(q)$ function. In reality, if a customer places an order of a size that is slightly less than a breakpoint, a fee associated with the corresponding breakpoint size will typically be charged. This results in the effective all-units quantity discount schedule which therefore must have flat portions at the breakpoint sizes. Figure 1 summarizes all types of price schedules discussed above.

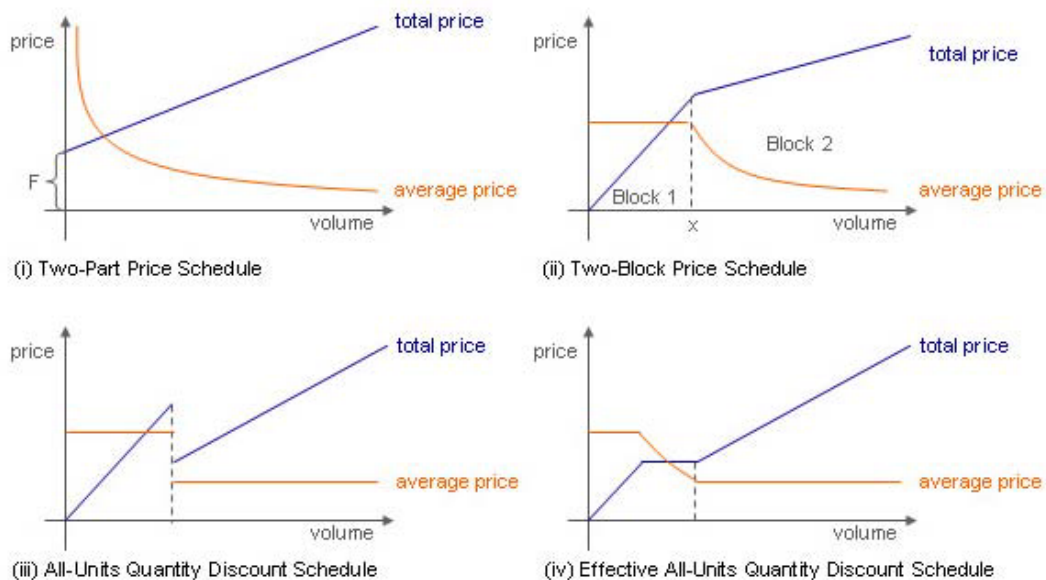


Figure 1: Price Schedule Types

2.3.2 Caps and Floors

Transaction fee caps and floors are commonly used instruments for securing minimum fees to cover costs and for giving buyers an incentive to place large orders. Since caps and floors are widely integrated into existing price schedules, we develop a formal representation analogue to the one in the previous paragraphs.

Floors are minimum fees and represent the lower border for the price schedule whereas caps are maximum fees representing the upper border. The first graph in Figure 2 on the left hand side illustrates the construction of a two-block schedule with both a cap and a floor. The graph on the right represents the resulting effective price schedule: The floor is represented by a fixed fee for quantities below x_c whereas the cap is essentially a fixed fee for quantities higher than x_c .

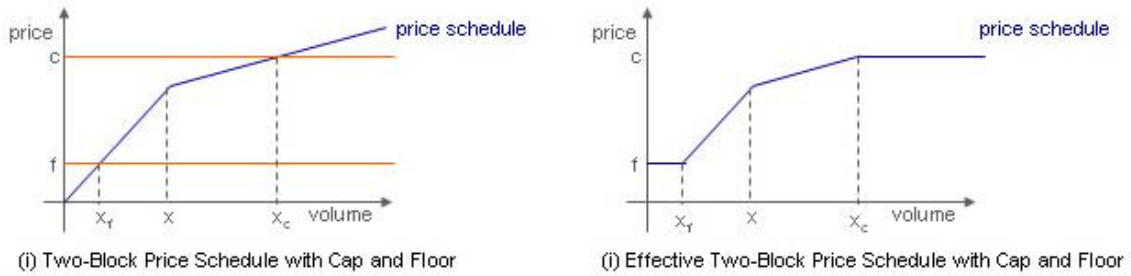


Figure 2: Graph of a two-block schedule with cap and floor

We can adapt our formal representation of a two-block schedule in a way so that it accommodates the construction of caps and floors. The breakpoint sizes x_f and x_c can be expressed as functions of the other parameters.

$$R(q) = \begin{cases} 0, & q = 0 \\ f, & 0 < q \leq x_f \\ p_1 q, & x_f < q \leq x \\ p_1 x + p_2 (q - x), & x < q \leq x_c \\ c, & q > x_c \end{cases} \quad \text{where} \quad \begin{aligned} x_f &= \frac{f}{p_1} \\ x_c &= \frac{c - (p_1 - p_2)x}{p_2} \end{aligned}$$

For marketing and communication purposes, companies usually communicate the block price schedule along with the values for the cap and the floor instead of integrating them into one combined effective price schedule. Thus, customers can easily identify minimum and maximum fees and make instant decisions based on these values. Nevertheless, for design and implementation purposes, a combined effective price schedule should be preferred.

2.3.3 Design Parameters

Several design issues are discussed in [Dola87], especially price schedule type, price schedule complexity, price schedule depth, and the qualifying unit base. Type refers to the question whether to use incremental or all units quantity discounts, complexity addresses the number of price breaks, depth the magnitude of the discounts, and qualifying unit base is the dependent variable which is used to determine whether a customer qualifies for a discount.

Concerning the price schedule type parameter, we can constrain design efforts to support n -block price schedules since a two-part schedule can easily be written as an equivalent two-block schedule with a prohibitively high marginal price in the first block effectively representing a fixed fee for the customers.

The price schedule complexity parameter has to be considered in a marketing context. Although it can be shown that $n+1$ breaks are strictly better than just n breaks [Moor84], it seems to be fair to assume that customers prefer a simpler schedule over a more complex one with a higher number of price breaks. Moreover, less complicated price schedules are cheaper to communicate and calculate.

Price schedule depth is an important economic design parameter because it has to be considered in the context of customer price elasticity of demand and willingness to pay expressed through the demand profile.

The qualifying unit base is a key parameter for the development of a price schedule framework including price schedules, rebates, caps, and floors. For example, the unit base could either be quantity, volume (in terms of monetary units), or average transaction fees paid. Which one of these variables are employed in the price or rebate schedule is a question whether it is needed to encourage high value transactions or influence the size of the aggregated transaction volume over a specified period of time.

2.4 Applications

2.4.1 New York Stock Exchange

Before the New York Stock Exchange introduced its new price schedule this summer, a multi-block schedule was in place that differentiated the transactions according to the number of shares. No charge was levied for system orders under 2,100 shares. For orders with a size between 2,100 and 5,000 shares, a transaction fee of \$0.0023 per share was charged. For order sizes above 5,000 shares, there was a charge of \$0.0001 for each additional share. In effect,

customers paid \$11.50 for the first 5,000 shares and \$0.0001 for every share which was above the breakpoint of 5,000. The total transaction fee was capped at \$80.00.

The NYSE then eliminated the free orders and employed a uniform marginal price of \$0.00025 per share capped at \$80.00 per transaction. Figure 3 illustrates the changes in the price schedule.

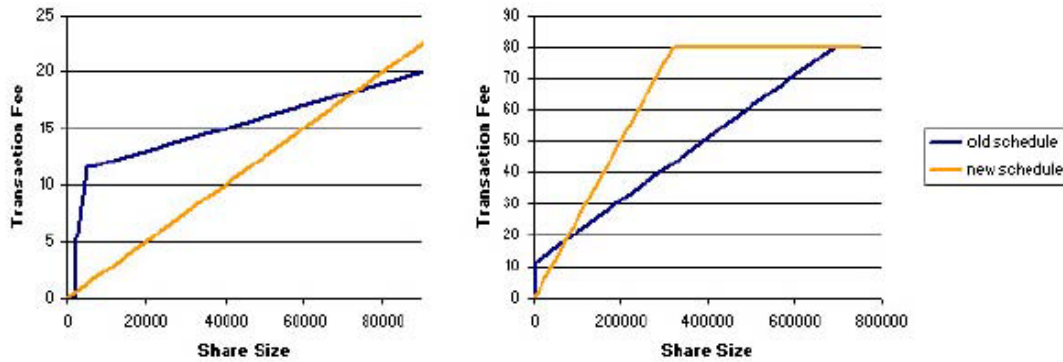


Figure 3: Comparison of the NYSE price schedule before and after the change. The graph on the left side is a magnification of the graph on the right side.

In addition to the \$80 cap per transaction, the NYSE has established a monthly cap of \$750,000 which means that companies do not pay more than \$750,000 per month. In effect, this constitutes an incremental rebate of 100% of the transaction fee since it applies to all future transactions after the qualifying amount has been reached.

2.4.2 NASDAQ Stock Market

The price schedule of the NASDAQ differs from that of the NYSE in a couple of points. Most importantly, the fully-electronic NASDAQ differentiates between removing and adding liquidity. Removing liquidity (i.e. placing a marketable order) for NYSE-listed securities is charged with a uniform price of \$0.0007 per share whereas adding liquidity (i.e. placing a non-marketable limit order) qualifies for a rebate of \$0.0005 per share executed. In addition, customers have to qualify for the rebate by reaching an average daily share volume of 5 million shares. If customers add more than 10 million shares in daily average, a rebate of \$0.0006 is given.

2.4.3 ebay

As another example of an electronic marketplace, the internet auction site ebay.com uses an incremental three-block price schedule for its execution fees which are based on the final closing price (which represents the quantity dimension in our formulation above). Price breaks

are at the \$25 and the \$1,000 level. The marginal prices in each block are decreasing and set at 5.25% in the range from \$0.01 to \$25.00, 3% in the range of \$25.01 to \$1,000.00, and 1.5% for the trailing block with final closing prices greater than \$1,000.00. Figure 4 illustrates the price schedule.

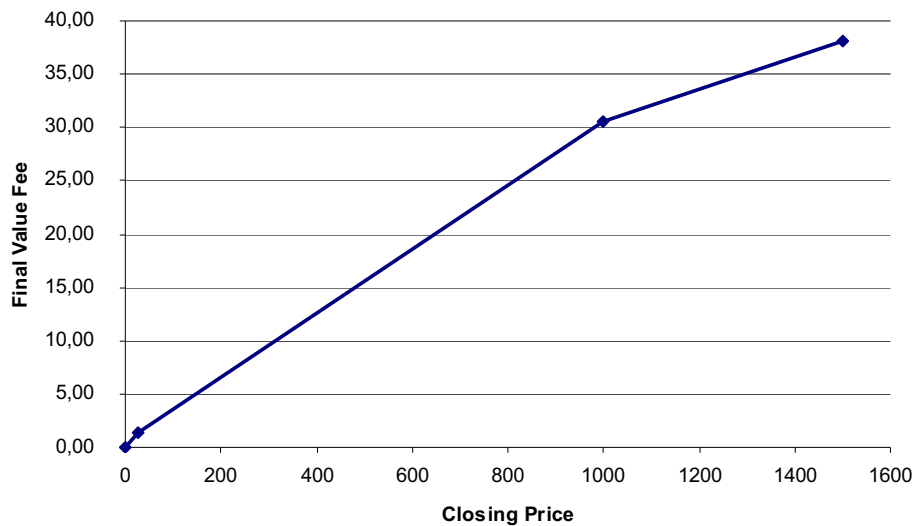


Figure 4: Ebay's Block Pricing Schedule

2.5 Section Summary

For the examples discussed in the previous section, the main motivation for nonlinear pricing is partial price discrimination. Due to the quantity discounts, the average price of small orders is higher than the average price of large orders. According to [Buch53], the underlying criterion for this motivation must be that the demand schedule of customers who place smaller orders is more inelastic over the relevant price range than that of customers who place large orders.

Why did the NYSE eliminate the price breaks at 2,100 and 5,000 shares and introduced a uniform price for the sizes of up to 320,000 shares instead? One explanation is that the NYSE wanted to simplify its price schedule for marketing purposes. Another explanation is that the NYSE took the chance to increase revenues when they had the opportunity. The price change allowed them to participate in the volume growth at the NYSE on the listed side of trading.

Such an effect on revenues must be considered in the light of customers' willingness to pay and customers' price elasticity of demand. Thinking in terms of price discrimination, altering the price schedule in such a way presumes that (i) customers with small orders up to 2,100 shares are price insensitive, (ii) the 5,000 shares price break has not proven to be useful for customer

segmentation, and (iii) customers with an order size of 73,000 to 690,000 shares are either price insensitive or they are expected to change their order behavior since they have to pay more with the new price schedule than with the old one.

The NASDAQ example illustrates how important it is for electronic exchanges to attract liquidity with an innovative price schedule. This price schedule relies on the fact that customers with urgent trading needs are less price sensitive than customers who do not need to trade immediately and thus place non-marketable orders.

The ebay.com price schedule illustrates the need for companies to employ price discrimination for profit maximization. It seems fair to assume that the majority of products offered on ebay is priced below the \$25 breakpoint, therefore average transaction fees are highest in this range in order to increase profits. Additionally, customers are assumed to be less price sensitive in this range than in ranges of higher closing price values.

To conclude, a structured approach for the design of price schedules is necessary in order to accomplish the different goals of companies in the execution business. Future work needs to be done on integrating price schedule design into the holistic market engineering process. Computer Aided Market Engineering [WeNH06] should be extended to include tools to analyze customer preferences and suggest optimal design parameters for price schedules based on economic as well as marketing principles.

3 Transaction Fee Experiment

3.1 Motivation

The previous sections have shown that information and knowledge about different customer types is necessary to design an optimal profit-maximizing price schedule. While empirical research on the distribution of transaction volume and frequency is needed to identify different customer segments, further research has to focus on the order behavior of customers when being confronted with changing price schedules.

In order to get an understanding about customer behavior following a change of the transaction fee schedule, a field experiment is conducted in which there are (a) different customer segments according to their initial endowment and (b) different transaction fees over the total trading period of three weeks. The control over the design of the transaction fee schedules makes it

possible to draw conclusions about the influence of transaction fees on the order behavior of investors.

3.2 Research Questions

There are several research questions that arise in the context of nonlinear transaction fee pricing.

- What is the price elasticity of demand in the context of transaction services? Are customers sensitive to price changes at all? Are there other – more important – determinants of their trading decision?
- Can investors be segmented into different groups that differ in their (a) price elasticity and (b) willingness-to-pay?
- How do investors react to changing price schedule types, i.e. do they take their distinctive features (e.g. price breaks in the stepwise price schedule) into account when placing their orders?
- How does trading volume and the number of transactions change when the transaction fee schedule changes?

Some of the questions above have been partly answered in related literature. [Epps76] develops a probability model which implies that the expected number of transactions per unit time is a decreasing linear function of the ratio of transaction cost to the security's price per share. Furthermore, the expected number of shares exchanged and the expected trading volume are decreasing functions of this ratio. He estimates these function for 20 stocks and finds that volume is indeed measurably responsive to changes in transaction costs. Furthermore, Epps estimates the demand for brokers' service to be -0.25. This estimation, however, is based on transaction records of the 1970s – today, people are likely to react very differently.

3.3 Experimental Design

We have conducted a trading experiment where participants could trade virtual stocks on an electronic platform within a time period of three weeks. During three weeks, participants faced transaction fees of different types. Order frequency and volume were measured and related to the price schedule in place.

3.3.1 Subjects and Groups

60 mostly undergraduate students in information engineering and management served as subjects for the experiment. They were grouped into 3 groups of traders: A, B, and C.

Group A was used as a benchmark group with a high initial endowment of shares and money and no transaction fees. Group B also had the high endowment but faced transaction fees during the experiment. Group C also faced transaction fees – the same as group B – but started with a much lower endowment, i.e. low income.

All groups were paid according to their trading performance. In order to encourage them to trade a certain amount of stocks, a weekly minimum trading volume requirement was in place which accounted for 10% of the overall income. If this requirement wasn't met in all three weeks, the payment was decreased to zero.

3.3.2 Transaction Fee Schedules

For groups B and C, a transaction fee schedule was in place that changed at the beginning of week two and three. Subjects were not told that (a) there was going to be a change of transaction fees, and (b) how that change would look like.

In the first week, the announcement of the transaction fee was:

- You are charged a fixed fee of 30 monetary units per executed order and a variable commission of 5% of the transaction volume. The minimum transaction fee, however, is 50 monetary units, the maximum fee 250 monetary units.

At the beginning of the second week, the following transaction fee schedule was announced:

up to a transaction volume of	monetary units	you are charged	monetary units
	500		50
	1500		100
	3000		150
	6000		200
from	6000	you are charged	250

The third week started with an announcement as follows:

- For each transaction, independent of the transaction volume, you are charged a fixed amount of 150 monetary units.

The different transaction fee schedules are summarized in the following figure:

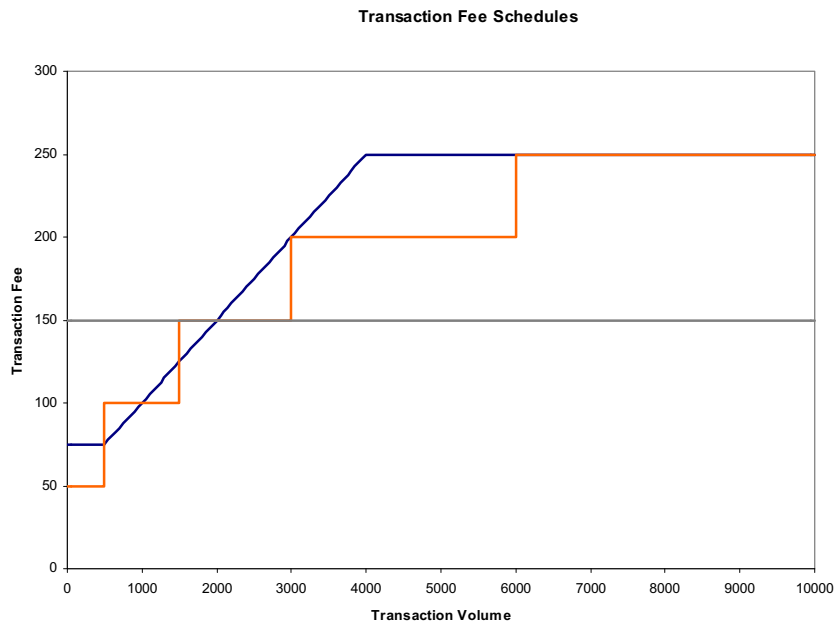


Figure 5: Transaction Fee Schedules for Weeks 1 to 3

3.3.3 Incentive Schemes

As mentioned above, subjects were all paid according to their performance. Since all participants of one group traded in the same continuous double auction, the game was a zero sum game where the amount gained by one part of the group has been lost by the other. Therefore, cheating wasn't attractive for the subjects. Since the transaction fees were subtracted from the payment, subjects had an incentive to minimize transaction fees in order to get a higher payment.

3.4 First (preliminary) results

We have focused on the average number of orders in each group. According to the changing fee schedules, the number of orders should have been decreasing, and the average transaction volume should have increased.

In the following graphs we have depicted box plots of the average number of orders in each group. We can see the order frequency (i.e. the actual number of orders per week) has decreased in groups B and C (where there were transaction fee schedules in place) whereas in group A, the average number of shares has remained constant (even with an increase from week 1 to week 2).

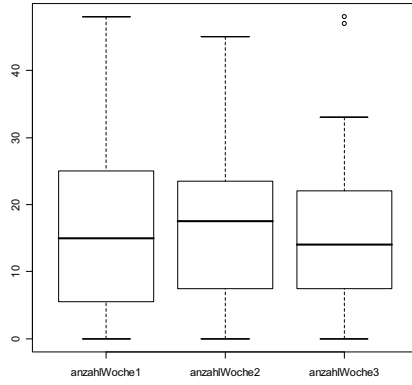


Figure 6: Average Number of Orders (Group A)

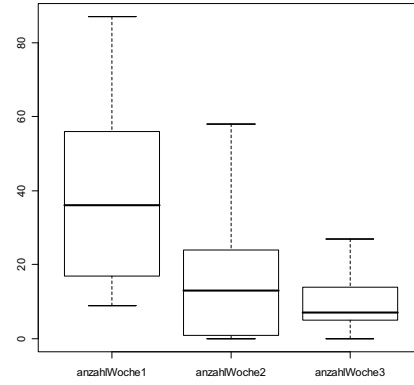


Figure 7: Average Number of Orders (Group B)

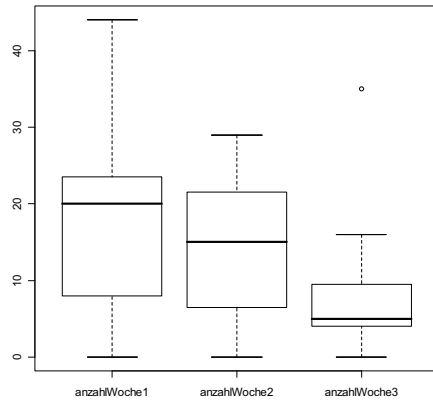


Figure 8: Average Number of Orders (Group C)

First results indicate that there is a significant increase in the average number of orders per period in the groups facing transaction fees whereas we do not find an increase in the number of orders in the test group without transaction fees.

4 Conclusion and Outlook

This paper has presented a classification of transaction fee schedules, a formal representation, and some real-world examples of nonlinear transaction fee schedules. In the second part of the paper we have outlined a field experiment where we investigated the effect of changing

transaction fees on the order behavior of subjects. While there are first promising results, there are a couple of points which still have to be addressed:

- Analytical solution: Based on the segmentation of customers, an analytical solution should be derived to determine the profit-maximizing price schedules. In particular, how many price breaks are possible, where are the price breaks depending on the segmentation, and what kind of price schedules are optimal.
- The objective of the stock exchange is not always profit-maximization. Some stock exchanges prefer to maximize customer welfare and act as not-for-profit organizations. In designing the optimal price schedule, this also has to be taken into account.
- Follow-up experiments should be designed to address questions which have not been known before. Based on the results of the experiment described, laboratory experiments should provide more insights into specific behaviour of the subjects.
- Real-world data from stock exchanges should be analyzed to verify the effect that changing transaction fees have on order behavior.

References

- [Buch53] Buchanan, J. M. "The Theory of Monopolistic Quantity Discounts," *The Review of Economic Studies*, (20:3), 1953, pp. 199-208.
- [Dola87] Dolan, R. J. "Quantity Discounts: Managerial Issues and Research Opportunities," *Marketing Science*, (6:1), 1987, pp. 1-22
- [Epps76] Epps, T. W., "The Demand for Brokers' Services: The Relation between Security Trading Volume and Transaction Cost," *The Bell Journal of Economics*, Vol. 7, No. 1. (Spring, 1976), pp. 163-194.
- [LeMe76] Leland, H. E., Meyer, R. A. "Monopoly Pricing Structures with Imperfect Discrimination," *The Bell Journal of Economics*, (7:2), 1976, pp. 449-462

- [Moor84] Moorthy, K.S., "Marketing Segmentation, Self-Selection, and Product Line Design," *Marketing Science*, (3:4), 1984, pp. 288-307
- [Skie99] Skiera, B., „Mengenbezogene Preisdifferenzierung bei Dienstleistungen“, *Beiträge zur betriebswirtschaftlichen Forschung* 90, Gabler, Wiesbaden, 1999
- [Simo82] Simon, H., "Preismanagement", Gabler, Wiesbaden, 1982
- [Tack89] Tacke, G., "Nichtlineare Preisbildung", *Beiträge zur betriebswirtschaftlichen Forschung* 64, Gabler, Wiesbaden, 1989
- [WeHN03] Weinhardt, C.; Holtmann, C.; Neumann, D.: Market Engineering, In: *Wirtschaftsinformatik* 45 (2003) 6, p. 635-640
- [WeNH06] Weinhardt, C., Neumann, D., Holtmann, C., "Computer-aided Market Engineering," *Communications of the ACM*, (49:7), 2006
- [Wils93] Wilson, R.B., *Nonlinear Pricing*, Oxford University Press, New York, Oxford, 1993

Outsourcing Risk Mitigation

A Study of Business Process Outsourcing arrangements in the German Banking Industry

Kim Wüllenweber

E-Finance Lab

Johann Wolfgang Goethe University

60054 Frankfurt am Main

wuellenweber@wiwi.uni-frankfurt.de

1 Introduction

Outsourcing entails a number of decisions regarding a variety of issues, be they the choice of the activities to outsource, the selection of a service provider, or the identification of the most appropriate way to manage an outsourcing contract. Managers hardly have all the information they need and have to decide on these issues under uncertainty. Yet, outsourcing research has provided a number of models and tools that are aimed at decreasing the level of uncertainty or that help in dealing with the uncertainty in which the decision is necessarily clouded. Such is the case for instance of the studies that examine the characteristics of the activities that are good candidates for outsourcing and of those that should be kept in-house; that compare outsourcing to insourcing, evaluating if, and how, in-house services could be reorganized to provide firms with benefits similar to those of outsourcing [LaHi95]; that attempt to identify those characteristics of a vendor-client partnership that will be conducive to outsourcing success [LeKi99] or that identify contractual means aimed at lowering or controlling uncertainty [AHP03;GeWü06].

This research contributes to outsourcing research by adopting a risk mitigation perspective in the delivery phase of an outsourcing project. Precisely, answers to the following statement being made in one of the best known reviews of outsourcing research are projected: “An interesting question would be under which circumstances different outcomes can be achieved using alternative governance structures.” [DGHJ04, p.80]. By analyzing the impact of different risk mitigation instruments - such as contract mechanisms, process standards or a good relationship with the service provider – on the achievement of outsourcing objectives, a

framework will be tested that helps reducing outsourcing risk and identifies management instruments that are most appropriate for the activities to be outsourced. As ‘activities’ business processes are chosen to close some gaps within the scarce knowledge of Business Process Outsourcing (BPO). Thus, this study aims at answering the following research question:

*What is the impact of risk mitigation instruments on the success of
Business Process Outsourcing arrangements?*

The following sections provide an overview how this research question will be approached.

2 Related Research

2.1 Outsourcing success

Multiple researches have been conducted to analyze outsourcing success. According to [DGHJ04], outsourcing success can be viewed either as the satisfaction with the processes/operations outsourced or as the realization of objectives.

For the purpose of this research, outsourcing success is achieved, if organization’s objectives are realized. Surprisingly, most studies using this approach measure success by the realization of benefits [LeKi99; RoCA01] but rarely compare or ‘weight’ these benefits to original objectives. The assumption within these studies is that the benefits realized are the same as the expected benefits before outsourcing. However, [GeYe06] showed that benefit’s perception and realization can be different. Thus, the focus is on differences in objective’s expectations and realization as measure of outsourcing success.

Outsourcing objectives have been analyzed in multiple outsourcing studies. As outlined by [LeKi99], the outsourcing objectives can be strategic, economic and technological. Strategic objectives refer to the bank’s ability to focus on core competencies. Economic objectives are associated with cost reductions through the utilization of economies of scale, scope and skill. Technological objectives refer to quality improvements, i.e. the bank’s ability to exploit modern IT technologies. By querying all of those objectives, [RoCA01] is followed to overcome critics of traditional uni-dimensional measurement.

Differences in objective’s expectations and realizations motivates for analysis of causes and drivers periling outsourcing success. Thus, “the assessment of risk associated with turning over

IS resources to third-party providers” [DGHJ04, p.87] is a promising research path and will be presented in the following section.

2.2 Outsourcing risk

In Rational Decision Theory, the concept of risk reflects the variation in the distribution of possible outcomes, their likelihood and their subjective values [Knig21]. This implies that alternatives are assessed on the basis of their expected probability distributions in conjunction with subjective values of the expected outcomes [MaSh87], which have to be precisely calculated [KaGa81]. Rational decision theory suggests that decision-makers deal with decisions under uncertainty in a rational way, i.e., by computing alternatives and selecting the option that best suits their personal risk-return profile, which is generally risk-averse. However, empirical studies indicate that this theoretical view is not consistent with how managers deal with risky choices in reality: several studies have shown that managers follow a less precise calculus, not using accurate probability calculations [MaSh87]. Thus, several researchers use a qualitative risk definition based on [BaRi05]: risk is a combination of *risk factors* and *undesirable outcomes*. Every risk factor causes a certain undesirable outcome; the risk factors represent the probability of undesirable outcomes. For example, if outsourced process is very specific, it will take the service provider a long time to understand the process. Meanwhile he will have a hard time to delivery the process in sufficient quality manners. Thus, process specificity (risk factor) causes service debasement (undesirable outcome). If process specificity is high, the probability of service debasement is also high.

In this research, this qualitative risk definition is used. The following sections provide for theoretical foundations of risk factors and undesirable outcomes.

2.2.1 Risk factors

As shown in other outsourcing studies (e.g. [BaRi3b]), Agency Theory (AT) and Transaction Cost Economies (TCE) can be used to identify risk factors. AT treats the difficulties that arise under conditions of incomplete and asymmetric information when a principal (the bank) hires an agent (the service provider) [Eise89]. From the principal’s perspective, these conditions entail continuous coordination and motivation issues. The general assumptions of agency theory are (1) the principal’s inability to observe the agent’s behaviour and (2) opportunism of both the principal and the agent. Opportunism can be explained using three concepts: moral hazard, adverse selection and imperfect commitment. *Moral hazard* emerges if the agent acts to the

disadvantage of the principal. This induces the risk of service debasement [BaRi3a]. The service provider might reduce his effort in collaborating with the bank and fail to improve the process if he could use his superior resources in other client engagements instead. *Adverse selection* is driven by the agent's allegation of superior capabilities and produces the problem of choosing the "wrong" agent. As outlined by [BaRi05], adverse selection emerges when the service provider does not have sufficient outsourcing experience, thus being unskilled in collaborating with the bank. *Imperfect commitment* refers to the reluctance of both the bank and the service provider to commit themselves. Thus, the bank and the service provider are likely to deny their promises and commitments. Typical occurrences of this behaviour are the refusal of activities or services due to unforeseeable events and ambiguities requirement definitions.

Transaction Cost Economics (TCE) allow to explore efficiencies in hybrid governance structures, such as joint ventures, strategic alliances or franchises [Will91]. Transactions are the transfer of goods or services between institutionally separable interfaces [Will91], which entail specific costs since TCE assumes limited rationality and opportunistic behaviour [DGHJ04]. Limited rationality refers to the limited capabilities of the bank that render "complete" contracts impossible. Opportunistic behaviour entails that this incapability could be exploited by the service provider which is the fundamental base of agency theory. TCE states that the more transactions (activities) occur (are performed), the more frictions and inefficiencies exist. These frictions and inefficiencies implicate additional costs, namely transaction costs. The magnitude of transaction costs is determined by three factors. Since outsourcing can be regarded as an organizational venture under uncertainty, these factors can be interpreted as factors increasing or decreasing risk. The factors are frequency of transactions, uncertainty of transactions and asset specificity: *Asset specificity* is the degree to which the assets needed to perform the activity are not transferable to other activities, i.e. are unique to certain activities. This raises the risk of lock-in so that is almost impossible to switch service providers or backsource. *Frequency* refers to the reoccurrence of the transaction or the asset used. If a transaction will often take place, it is regarded as frequent. However, if the transfer of activities occurs rarely, the cost of setting-up governance mechanisms for this transaction seem too high. Thus, if there is a low frequency of transaction, risks arise. *Uncertainty* is segregated in environmental and behavioural uncertainty. Environmental can be both the environment within the company (i.e. other departments) and forces endogenous of the firm. As outlined by [LaHi95], environmental uncertainty can also be separated in business driven uncertainty and technology driven

uncertainty. Business driven uncertainty describes possible changes due to shifting business developments. These new developments might result in new requirements to be implemented in the process. Technology driven uncertainty refers to possible alteration of required technical skills. If a new technology is used in an existing application or a new application uses a well-established technology, ambiguity about process continuity might arise. Behavioural uncertainty refers to unpredictable behaviour of those completing the tasks. In particular, appropriate assessment of the vendor performance might be difficult [PoZe02]. However, not being able to assess vendor performance does not necessarily entail that opportunistic behaviour is present. But if it is, behavioural uncertainty becomes even more severe.

AT and TCE can be used to explain risk factors. The following section elaborates on undesirable outcomes as the second concept of the risk definition.

2.2.2 Undesirable outcomes

The risk factors from section 2.2.1 cause different undesirable outcomes. As outlined by [BaRi3b], these outcomes can be cost (increase cost of services, costly contract amendments, etc.) or performance (e.g. service debasement) related. [GeWW06] extends this risk composition by strategic (e.g. loss of organizational competence) and psycho-social risks (e.g. loss of personal reputation). As this research focuses on risks inherent to outsourcing arrangements, financial, performance and strategic related undesirable outcomes are analyzed. Psycho-social concerns are excluded that only address risks inherent to manager's career path. The analysis of undesirable outcomes will enable us to have a precise picture why objective's expectations and realization diverge in outsourcing arrangements. However, this divergence only emerges in ventures where risk was not properly managed. A review of management instruments intended to successfully tackle risk is needed. Thus, the following section presents academic findings on risk mitigation in outsourcing context.

2.3 Outsourcing risk mitigation

The objective of risk mitigation is to reduce the level of risk exposure of a given business venture [AuBR05]. Given an expected level of benefits from an outsourcing strategy, a rational, profit maximizing, and risk-averse decision maker wishes to minimize the risk exposure of the project or the strategy. A boundedly rational decision maker may wish, for the same given level of benefits, to bring the level of risk exposure below some threshold or acceptable level.

Reviewing outsourcing literature reveals that most research on risk mitigation is done by analyzing the outsourcing contract (see e.g. [AHP03]). [AHP03], [AuBR05] and [GeWü06] showed that detailed and precisely service levels, penalties for underperformance, bonus for outstanding performance, stipulated cost reductions over time and renegotiation options help to reduce outsourcing risks. To effectively use contractual provisions, the outsourcer needs to have resources and capabilities to monitor service quality and enforce service violations. Thus, the governance or management structure has mitigating influence on outsourcing risks [LeMK04]. This includes not only managerial experience and support from ‘upper echelon’ but also capabilities to choose outsourcing strategies that profile for a low risk profile upfront – such as selective sourcing [LaWF96; CuWi98; WiLK99; BaRi3b].

However, the contract and governance structure might not reduce all outsourcing risks. If measurement problems exist [BaRi3b] and/or asset specificity is high [BeRi03], a good relationship with the service provider might prevent from opportunistic behaviour and the problem of writing a ‘complete’ contract [PoZe02; MaBW06]. The so-called ‘psychological contracts’ entail non-written obligations of the service provider and serves as a complement to contractual written obligations [PoZe02].

2.4 Concepts of ‘fit’

Risk, risk mitigation and success are multi-dimensional constructs. Researchers in software engineering have acknowledged the problem of their relation (e.g. [BaRT01]): why can low risky projects be unsuccessful? Do projects with low risk need any risk mitigation? Do risk mitigation instruments impact on risk factors, on the impact of risk factors on undesirable outcomes, undesirable outcomes or the impact of undesirable outcomes on success? To account for this problem, an integrative perspective is adopted as suggested by [BaRT01]. The concepts of ‘fit’ from the contingency perspective research stream in Organization Theory will be adopted: it is said that the conventional treatment and use of control mechanisms in outsourcing ventures might not be optimal for dealing with the agility and new cultural dynamics of the outsourcing environment today [GoKR04]. As simple outsourcing evolves into more complex models of IT outsourcing deals, unseen factors such as trust and culture become more important. Thus, it is generally expected that management mechanisms require even more sophisticated contracts that address additional complexity as ambitions for outsourcing encompass multiple objectives.

When adopting a contingency approach, researchers must carefully define their conceptualization of fit [Venk89]. Definitional rigor is critical, since different conceptual

definitions of fit imply different meanings of a contingency theory and different expected empirical results. [Venk89] proposed a framework that defined six different perspectives from which fit could be studied. These six perspectives can be classified into two broad categories, according to the number of variables being simultaneously examined: the 'reductionist' conceptualization of fit and the 'holistic' conceptualization [BaRT01]. The 'reductionist' conceptualization "is based on a central assumption that the co-alignment between two constructs can be best understood by analyzing pair wise co-alignment among individual dimensions of those constructs. Three of the six perspectives of fit defined by [Venk89] belong to this conceptualization. They are the moderation, the mediation, and the matching perspectives. The holistic conceptualization of fit emphasizes the systemic nature of co-alignment, wherein several variables are examined simultaneously. Fit as co-variation, fit as gestalt, and fit as profile deviation all pertain to the holistic perspective. Several researchers have argued that the latter conceptualization was richer, because of its ability to retain the complex and interrelated nature of the relationships between variables [DrVe85; Venk89].

The conceptualization of risk mitigation in outsourcing context comprises a variety of risk mitigation instruments (see section 2.3.) which can reduce the severity of risk factors and undesirable outcomes simultaneously. Thus, a holistic conceptualization of fit is suggested. The framework of [Venk89] further characterizes each perspective along two other dimensions: the presence of a criterion variable, and the degree of specificity of the functional form of fit. The degree of specificity expresses how good the interplay of the variables is understood in literature. The criterion variable is considered as a benchmark for a good fit; the functional form of fit describes how the variables mathematically relate. In this study, there is a criterion variable – outsourcing success – but the degree of specificity of the relationship studied is low. For example, although the literature provides recommendations for contract clauses in case of high-risk outsourcing ventures, it says little how to write a contract for low-risk outsourcing ventures. Thus, the relation between the constructs of risk mitigation and success is unclear resulting in a low degree of specificity. A low degree of specificity suggests that - among the three perspectives which belong to the holistic conceptualization – fit as profile deviation best reflects the characteristics of the contingency phenomenon studied here. Consequently, fit is defined here as the degree of proximity to an externally specified profile or pattern, whereby "patterns of consistency among dimensions of organizational context, structure, and performance" [DrVe85, p. 520] are simultaneously examined. Adapting Venkatraman's

definition of fit as profile deviation to the context of outsourcing projects, this approach implies that if an ideal pattern for a risk mitigation profile is specified for a particular level of risk exposure, an outsourcing project's degree of adherence to such a multidimensional profile will be positively related to success. Conversely, deviation from this profile would imply a low degree of fit, resulting in a negative effect on success.

According to Venkatraman, "this perspective allows a researcher to specify an ideal profile and to demonstrate that adherence to such a profile has systematic implications for success. An important issue when adopting this conceptualization of fit is the development of an ideal pattern or profile. According to Venkatraman, there are two obvious alternatives: (a) developing them theoretically, or (b) developing them empirically. Venkatraman argues that while the first alternative is "intuitively appealing, the operational task of developing such a *qualitative* profile with *numerical* scores along a set of dimensions is difficult" [Venk89, p. 434]. Moreover, outsourcing research does not provide sufficiently precise levels for the different risk mitigation approaches it recommends for different levels of risk. Consequently, developing theoretically ideal profiles based on past research is hardly feasible. The second alternative is to develop an ideal profile using a calibration sample - generally defined as the data points that are closer to the top of the performance scale. Then a measure of the closeness between the two patterns (the Euclidean distance between the ideal pattern for an outsourcing venture and venture's actual pattern) is correlated with indicators of outsourcing success. Significant - but negative - correlations provide evidence for the presence of contingent relationships, as defined by the profile deviation perspective of fit.

The concepts of 'fit' have been used in two well-known outsourcing studies before [GoKR04; LeMK04]. However, these studies have not taken the perspective of risk and risk mitigation. The only available study where risk and risk management have been analyzed with the concepts of 'fit' has been done by [BaRT01]. Yet, the authors did not analyze outsourcing arrangements but (internal) software development projects. Therefore, the approach of [BaRT01] is taken and transferred to BPO. To transfer all concepts of this study to BPO context, the following section provides insights from literature on BPO risks, BPO risk mitigation and BPO success.

2.5 Business Process Outsourcing

BPO has attracted researches and practitioners and is acknowledged to be one of the largest areas of growth in the outsourcing market [Gart04]. For the purpose of this paper, BPO is defined as the delegation of one or more entire business processes to third party providers,

including the software and hardware that support those processes [HaMe00]. A business process is defined as a “set of logically related tasks performed to achieve a defined business outcome” [DaSh90, p. 12]. Thus, BPO is the combination of application development/maintenance, IT infrastructure outsourcing and the outsourcing of business activities which are not IT supported such as business process re-design.

Studies on BPO risk, risk management and success are rare. One of the most comprehensive analyses is survey of [GeWW05]. The researches took a perceptual perspective on BPO risks and BPO benefits and surveyed managers in charge of back-office processes in German banks. Those managers were asked about their general perception of BPO risks and benefits without considering their own process. Even those managers who have outsourced their process were not asked about the risks and benefits of their outsourcing projects. The study provides for (at least) three interesting findings:

- 1.) managers perceive high risk with outsourcing business processes;
- 2.) strategic risks and benefits play a very important role when considering BPO;
- 3.) quality benefits are more important on the benefit side than quality risks on the risk side.

[GeWü06] conducted a series of case studies to identify mitigation instruments who effectively counter risk. The authors have identified several contractual risk mitigation instruments but did not focus on relational and structural (such as capabilities of the retained organization) aspects of risk mitigation. As ‘effective’ contractual risk mitigation instruments the authors identified detailed and precisely defined service levels, performance based payment, liability clauses and renegotiation options.

[RoCo04] outline that BPO are beneficial if the processes are not new to the (service provider) market. In addition, management structure and precise contractual obligations of the service provider should overcome the problems of a relatively new outsourcing phenomenon as BPO. As the cost structure of a business process is sometimes hard to identify, great emphasis should be taken to determine desired cost savings before outsourcing.

3 Research Model and Methodology

3.1 Research model

The research model is presented below. The main hypothesis (H1) is that a good fit between outsourcing risk and outsourcing risk mitigation is positively related to outsourcing success. This hypothesis has been tested by [BaRT01] surveying software development projects who have not been outsourced. Transferring the model to BPO context, it has to be considered that outsourcing projects have a different risk profile, necessitate different risk mitigation instrument as external service provider are involved and strive for other benefits as success is much more long-term oriented.

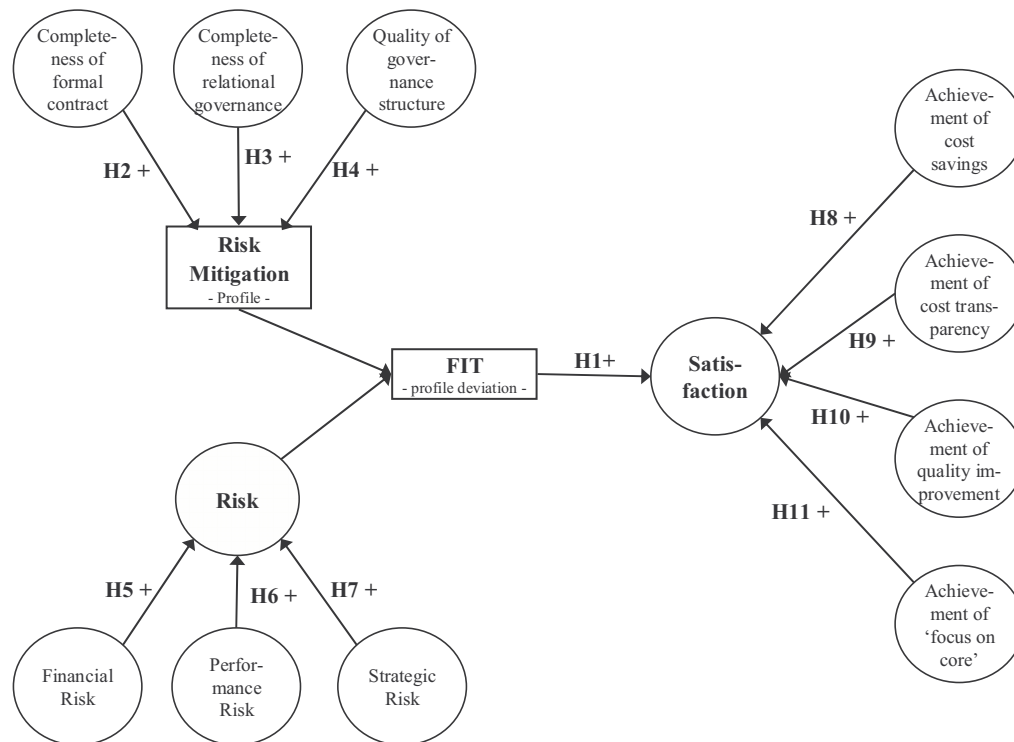


Figure 1 - Research Model

3.1.1 Formation of outsourcing risk mitigation

As previously outlined, three variables forming the ability of mitigating or controlling risk will be analyzed: (a) the outsourcing contract, (b) trust or a good relationship between the outsourcer and the service provider, and (c) governance structure of the outsourcing project.

The importance of the contract has been stressed by several researchers. It regulates the venture. If contract and negotiation processes are thoroughly defined, a company has the necessary protection for dealing with daily actions and results, for settling disputes or for terminating the relationship. The contract is regarded as the most important instrument to mitigate outsourcing risk [AuBR05]. In particular, precisely defined and detailed Service Level Agreements (SLAs), a penalty-reward-system, renegotiation options, and liability clauses help mitigating outsourcing risks [GeWü06].

No.	Description
H2	The higher the completeness of the formal contract, the lower the outsourcing risks.

Table 1 - Hypothesis on formal contract completeness

Despite the importance of a written formal contract, outsourcing research has acknowledged that legal contracts alone are insufficient, given the complexities of real-life outsourcing arrangements and the rapid changes in technology and organizational environments [KoAS04]. Specifying long-term exchange relations is complex as they have to cover various unspecified obligations and thus are inherently incomplete. Incomplete contract theory suggests that not all possible events during an outsourcing venture can be predicted and included upfront in contract clauses [BeWi03].

Thus, understanding the relationship that arises in an IT outsourcing arrangement is crucial, since it comes about not only through the operationalization of the legal contract, but also as a natural consequence of the resulting dependency between the two parties engaged in the outsourcing arrangement [KeWi00]. A good relationship with the service provider might complement contractual means and, thus, can be considered as risk mitigation instrument. Although relational aspects hardly influence process-specific risks – such as undesirable outcomes from asset specificity –, a good relationship can be used as a social or psychological contract inducing non-written obligations by the service provider [Ross95; KeWi00; BaRi3b; KoAS04]. Early theories on contracts posit that inter-organizational relationships such as outsourcing are in reality governed by more than legal exchanges [Macn80]. Macneil thus proposes revamping contract theory so that it caters for prior and future actions of individuals participating in an outsourcing relationship [KeWi00].

No.	Description
H3	The higher the completeness of relational governance, the lower the outsourcing risks.

Table 2 - Hypothesis on psychological contract completeness

In addition to contract and relationship related instruments, the governance structure of the outsourcing venture has a positive (mitigating) effect on the risk profile. The governance structure is the result of the way an organization sets up or shapes the outsourcing arrangement in accordance with business needs from the beginning [LeMMK04]. Thus, governance structure is conceptualized as all management-specific characteristics providing for a low risk profile upfront – such as outsourcing management experience [BaRi3b], selective sourcing [LaWF96], acquisition of external expertise [BaRi3b] or top management support [WiLa99].

No.	Description
H4	The better outsourcer's internal governance structure, the lower the outsourcing risks.

Table 3 - Hypothesis on quality of governance structure

3.1.2 Formation of outsourcing risk

As outlined by [GeWW06], BPO risk is formed by financial, performance, strategic risk and psycho-social risk. The authors did not ask for 'real' risk of outsourcing ventures. Instead they advocate a perceived risk lenses based on individual risk perception. As this research addresses the risks of the outsourcing arrangement and not the manager in charge, the construct of psycho-social risk is excluded. In addition, only managers who have outsourced their process are queried. Therefore, this research focuses on financial, performance and strategic risks which are present or actual (instead of perceptual in the study of [GeWW06]) during the outsourcing venture. Nevertheless, the hypotheses developed by [GeWW06] can be used for the study:

No.	Description
H5	The higher the financial risks of a BPO venture, the higher the overall risk of a BPO venture.
H6	The higher the performance risks of a BPO venture, the higher the overall risk of a BPO venture.
H7	The higher the strategic risks of a BPO venture, the higher the overall risk of a BPO venture.

Table 4 - Hypotheses on BPO risks

3.1.3 Formation of satisfaction

There are several studies on outsourcing satisfaction or outsourcing success (see [DGHJ04] for an overview). This research adopts the BPO benefits formation developed by [WGFW05]. However, the constructs will be measured differently as the focus is whether benefits are achieved.

No.	Description
H8	The higher the cost savings achieved, the higher the outsourcing success.
H9	The higher the cost transparency achieved by outsourcing, the higher the outsourcing success.
H10	The higher the quality improvements achieved, the higher the outsourcing success.
H11	The more the outsourcer has been able to focus on its core competencies, the higher the outsourcing success.

Table 5 - Hypotheses on BPO success

3.2 Methodology

As the research object for this study managers responsible for one of four designated business processes in Germany's 500 largest banks are chosen. These processes are generally not regarded as areas of core competence for banks [LaPö04]: back office/settlement processes for transactions in securities, consumer credits, credit cards and domestic payments. All 500 top banks will be contacted by phone to personally identify the managers and the sourcing status of the process - that is whether the process is outsourced or provided internally. As the research concentrates on outsourced processes, (bank-) internally provided processes are excluded from the sample. The remaining sample will be the basic population and receive the questionnaire.

The telephone activities started on September 1st and will last until October 15th. The deadline for answering the questionnaire will be November 15th. Managers not returning the questionnaire will receive a phone call asking if they needed assistance. Deadline for latest responses will be December 22nd.

3.3 Limitations of the study

There are several limitations of the study which have to be considered upfront. Firstly, there is a time issue. Instruments are projected which provide for a low risk profile and guarantee the success of an outsourcing venture. However, there will be no opportunity to ask when those instruments have been implemented. Thus, it cannot be determined if the effects of this instrument have already occurred. The same applies to the concept of success. If the outsourcing project started 'a couple of weeks ago', objective's achievement will hardly be determinable. To control for the time issue, it will be asked when the process has been sourced out. Nevertheless, it won't be possible to ask for the implementation date for every single risk mitigation instrument.

Secondly, the object in question is the outsourcing arrangement. There is confidence that the manager in charge will be knowledgeable about all risk, risk management and success related

matters. However, the manager might not be certain with all the details, e.g. individual contract clauses or the risk attitude of other managers. To overcome this limitation, more abstract questions will be used without periling the potential results.

A third point considers the focus of the relevant target group of the study. Although research in outsourcing management has increased the collective understanding of e.g. success factors, risk issues, or the value of a legal contract, the vast majority of studies – including ours – adopt only the customer perspective. Studies that incorporate both the customer (bank) and the vendor perspective remain still rare [KoAS04]. However, as outsourcing arrangements incorporate mutual agreements and obligations both perspectives should be addressed to gain a comprehensive understanding [KoAS04]. In this study, the customer perspective is adopted for a specific reason: the interest is in risk mitigation approaches that can be influenced by the customer. Thus, the focus is issues that are perceived or initiated by the customer, not by the vendor. Moreover, since the overall perception of e.g. relationship issues will be identified, a comprehensive view of the overall outsourcing arrangement is taken.

References

- [AuBR05] Aubert, B. A., B. Bahli and S. Rivard (2005). "A Framework for Information Technology Outsourcing Risk Management." *The DATA BASE for Advances in Information Systems* 36(4).
- [AHPR03] Aubert, B. A., J.-F. Houde, M. Patry and S. Rivard (2003). *Characteristics of IT Outsourcing Contracts*. 36th Hawaii International Conference on System Sciences, Hawaii.
- [BaRi3a] Bahli, B. and S. Rivard (2003a). *A Validation of Measures associated with the Risk Factors in Information Technology Outsourcing*. 36th Hawaii International Conference on System Sciences.
- [BaRi3b] Bahli, B. and S. Rivard (2003b). "The information technology outsourcing risk: a transaction cost and agency theory-based perspective." *Journal of Information Technology* 18(September): 211–221.

- [BaRi05] Bahli, B. and S. Rivard (2005). "Validating measures of information technology outsourcing risk factors." *Omega* 33: 175-187.
- [BaRT91] Barki, H., S. Rivard and J. Talbot (2001). "An integrative contingency model of software project risk management." *Journal of Management Information Systems* 17(4): 37-69.
- [BeWi03] Beulen, E. and P. Ribbers (2003). *IT Outsourcing Contracts: Practical Implications of the Incomplete Contract Theory*. 36th Hawaii International Conference on System Sciences, Hawaii, USA.
- [CuWi98] Currie, W. and L. Willcocks (1998). "Analysing four types of IT sourcing decisions in the context of scale, client/supplier interdependency and risk mitigation." *Information Systems Journal* 8: 119-143.
- [DaSh90] Davenport, T. H. and J. E. Short (1990). "The New Industrial Engineering: Information Technology and Business Process Redesign." *Sloan Management Review* 31(4): 11-27.
- [DGHJ04] Dibbern, J., T. Goles, R. Hirschheim and B. Jayatilaka (2004). "Information Systems Outsourcing: A Survey and Analysis of the Literature." *The DATA BASE for Advances in Information Systems* 35(4): 6-102.
- [DrVe85] Drazin, R. and A. H. Van de Ven (1985). "Alternative Forms of Fit in Contingency Theory." *Administrative Science Quarterly* 30(4): 514-539.
- [Eise89] Eisenhardt, K. M. (1989). "Agency Theory: An Assessment and Review." *Academy of Management Review* 14(1): 57-74.
- [Gart04] Gartner (2004). *Outsourcing Market View, What the Future Holds*. Gartner Dataquest.
- [GeWü06] Gellings, C. and K. Wüllenweber (2006). *The Impact of Contract Design on Outsourcing Success: Case Studies on Risk Mitigation Measures*. 12th Americas Conference on Information Systems, Acapulco, Mexico.

- [GeWW06] Gewald, H., K. Wüllenweber and T. Weitzel (2006). "The Influence of Perceived Risks on Banking Managers' Intention to Outsource Business Processes - A Study of the German Banking and finance Industry." *Journal of Electronic Commerce Research* 7(2): 78-96.
- [GeYe06] Gewald, H. and P. Yetton (2006). Differences in Benefit Perception According to Alternative Statuses of Business Process Outsourcing Adoption. Twelfth Americas Conference on Information Systems, Acapulco, Mexico.
- [GoKR04] Goo, J., R. Kishore and H. R. Rao (2004). Managing IT Outsourcing Relationships Using Service Level Agreements (SLAs): A Multi-Dimensional Fit Approach. Tenth Americas Conference on Information Systems, New York.
- [HaMe00] Halvey, J. K. and B. M. Melby (2000). *Business Process Outsourcing - Process, Strategies and Contracts*. New York, John Wiley & Sons.
- [KaGa81] Kaplan, S. and B. J. Garrick (1981). "On The Quantitative Definition of Risk." *Risk Analysis* 1(1): 11-27.
- [KeWi00] Kern, T. and L. P. Willcocks (2000). "Exploring information technology outsourcing relationships: theory and practice." *Journal of Strategic Information Systems* 9: 321 - 350.
- [Knig21] Knight, F. H. (1921). *Risk, Uncertainty and Profit*. Chicago.
- [KoAS04] Koh, C., S. Ang and D. W. Straub (2004). "IT Outsourcing Success: A Psychological Contract Perspective." *Information Systems Research* 15(4): 356–373.
- [LaHi95] Lacity, M. C. and R. Hirschheim (1995). *Beyond the Information Systems Bandwagon: The Insourcing Response*. Chichester, John Wiley & Son.
- [LaWF96] Lacity, M. C., L. P. Willcocks and D. F. Feeny (1996). "The Value of Selective IT Sourcing." *Sloan Management Review* Spring: 13-25.

- [LaPö04] Lamberti, H.-J. and A. Pöhler (2004). Die Industrialisierung des Backoffices am Beispiel der etb. Management von Transaktionsbanken. H.-J. Lamberti, A. Marliere and A. Pöhler. Berlin, Spronger: 3-38.
- [LeKi99] Lee, J.-N. and Y.-G. Kim (1999). "Effect of Partnership Quality on IS Outsourcing Success: Conceptual Framework and Empirical Validation." *Journal of Management Information Systems* 15(4): 29-61.
- [LeMK04] Lee, J. N., S. Miranda and Y.-M. Kim (2004). "IT Outsourcing Strategies: Universalistic, Contingency, and Configurational Explanations of Success." *Information Systems Research* 15(2): 110–131.
- [Macn80] Macneil, I. R. (1980). *The New Social Contract: An Inquiry into Modern Contractual Relations*. New Haven, CT, Yale University Press.
- [MaBW06] Mani, D., A. Barua and A. B. Whinston (2006). "Successfully Governing Business Process Outsourcing Relationships." *MIS Quarterly Executive* 5(1): 15-29.
- [MaSh87] March, J. G. and Z. Shapira (1987). "Managerial Perspectives on Risk and Risk Taking." *Management Science* 33(11): 1404-1418.
- [PoZe02] Poppo, L. and T. Zenger (2002). "Do Formal Contracts and Relational Governance Function as Substitutes or Complements?" *Strategic Management Journal* 23(8): 707-725.
- [RoCo04] Rouse, A. C. and B. Corbitt (2004). IT-supported business process outsourcing (BPO): The good, the bad and the ugly. Eight Pacific Asia Conference on Information Systems, Shanghai, China.
- [RoCA01] Rouse, A. C., B. J. Corbitt and B. A. Aubert (2001). Perspectives on IT-Outsourcing Success: Covariance Structure Modelling of a Survey of Outsourcing in Australia. Ninth European Conference on Information Systems, Bled, Slovenia.

- [Ross95] Rousseau, D. M. (1995). *Psychological Contracts in Organizations: Understanding Written and Unwritten Agreements*, Sage Publications, Inc., Thousand Oaks, CA.
- [Venk89] Venkatraman, N. (1989). "The concept of fit in strategy research: toward verbal and statistical correspondence." *Academy of Management Review* 14(3): 423-444.
- [WiLa99] Willcocks, L. P. and M. C. Lacity (1999). "IT outsourcing in insurance services: risk creative contracting and business advantage." *Information Systems Journal* 9: 163-180.
- [WiLK99] Willcocks, L. P., M. C. Lacity and T. Kern (1999). "Risk Mitigation in IT outsourcing strategy revisited: longitudinal case research at LISA." *Journal of Strategic Information Systems* 8(3): 285-314.
- [Will91] Williamson, O. E. (1991). "Comparative Economic Organization: The Analysis of Discrete Structural Alternatives." *Administrative Science Quarterly* 36(2): 269--296.
- [WGFW05] Wüllenweber, K., H. Gewalt, J. Franke, T. Weitzel and W. König (2005). *Auslagerung von Geschäftsprozessen - Nutzen und Risiken*. Frankfurt am Main.

Conception of XBRL Use in the Financial Reporting Supply Chain

Exposé for the Doctoral Thesis

Maciej Piechocki

Fakultät für Wirtschaftswissenschaften
Technische Universität Bergakademie Freiberg

09599 Freiberg

maciej.piechocki@bwl.tu-freiberg.de

Supervisor: Univ.-Prof. Dr. Carsten Felden

Fakultät für Wirtschaftswissenschaften
Technische Universität Bergakademie Freiberg

09599 Freiberg

Planned submission: April 2007

Abstract

The introduction of the International Financial Reporting Standards and Basel II in European Union or the Sarbanes-Oxley Act in the United States influenced the domain of financial reporting. Together with the harmonisation and standardisation efforts from the domain perspective an attempt was undertaken to standardise the technical aspects of financial reporting. With means of XBRL the term *Financial Reporting Supply Chain (FRSC)* found broader use, implying the flow of the financial information from the business operations, through internal and external reporting, to the end user of the financial report. However the *FRSC* acts more as a consulting term and can be rarely found in the literature. In order to analyse the impact of XBRL on the *FRSC*, the *FRSC* itself needs to be analysed and formalised. First then the consequences of XBRL use within the *FRSC* can be assessed and measured. This exposé should reflect the methodology and first results of the literature review included in the doctoral thesis. The exposé is constructed consequently to the doctoral thesis presenting its major conclusions.

1 Introduction

The financial reporting world has faced a number of changes in recent years. The Internet with XML technologies and especially eXtensible Business Reporting Language (XBRL) technology has impacted what is recognised as the *Financial Reporting Supply Chain (FRSC)* [RoSt2006, 532]. Some claims in the market report XBRL to reduce inefficiencies, automate and optimise the *FRSC* [Hann2004, 55; RoSt2006, 530; NuSt2002, 457]. Nevertheless the real nature of the impact remains unclear [Sutt2006, 3]. The growing number of XBRL projects around the world and strong interest from bodies such as Security Exchange Commission in the United States (SEC), Central European Banking Supervisors in the European Union (CEBS) or International Accounting Standards Board (IASB) building XBRL taxonomies demonstrate the need for research in the area of XBRL application in the context of financial accounting and Accounting Information Systems (AIS) as well as in the *FRSC* context.

The first goal of the doctoral thesis is to properly define and set the boundaries of the *FRSC*. It is done by modelling the inter-organisational financial reporting domain. The second goal is to critically assess the impact of XBRL on the *FRSC*. The assessment is conducted by enhancing the *FRSC* model with XBRL technologies creating *XBRL FRSC* model and verifying the model with a prototype. The doctoral thesis consists of six chapters. The introductory chapter introduces the hypothesis as well as the research methodology. The next two chapters *Domain Analysis* and *Technical Analysis* are based on literature review of financial reporting and XBRL domains. The main part of the doctoral thesis comprises the conceptual model of the *FRSC*, its enhancement with the XBRL technology in the fourth chapter and the experimental measures in the fifth chapter. The doctoral thesis outlines the research results and conclusions in the final chapter. In the further part of the exposé the detailed description of individual chapters is provided.

The first introductory chapter gives an overview of the problem that exists in the reporting domain. The current business and especially financial reporting processes and information flows cannot be fully automated, are often not standardised and identified as inefficient [RoSt2006, 530; GeSH2005, 607]. The extant research is focused on the AIS aspects [RaKO2006; RoSt2006; GeSH2005; BaSS2005] and rarely deals with the inter-organisational reporting aspects [Teix2005, 2; Sutt2006, 4]. An interesting research in the area of database supported pub-

lication of financial reports provides Gassen [Gass2000, 164] but the XBRL technology is immature at the point of his research to consider it in the context of the possible reporting medium. The scope of the doctoral thesis is to research the inter-organisational reporting supply chain. The thesis addresses aspects of accounting processes, report preparation processes and reporting processes with a special focus on the information and document flow as well as participants of the processes. The thesis analyses the financial reporting for non-financial institutions, especially for listed companies. But most of the applications are also important for the small and medium enterprises (SMEs). The thesis analyses the financial and tax audit domain only in the context of the information needs and not the detailed audit processes. Although the reporting process of a subsidiary to its parent company is analysed the consolidation processes as such are out of scope of the thesis. Also the accounting processes described in the second chapter are analysed only in context of the creation of financial reports and not accounting of various transactions in this context.

The main hypotheses are stated as follows the Financial Reporting Supply Chain can be identified and modelled within and between organisations with the second hypothesis communicating that the XBRL Financial Reporting Supply Chain model can be implemented on an inter-organisational basis. While the first hypothesis is the basis for further evaluations and is not subdivided the second one consists of a group of hypothesis. Together with the construction of the XBRL FRSC model the following statement is verified XBRL increases the efficiency and reduces the costs of the Financial Reporting Supply Chain for both sending and receiving institutions. Together with XBRL use along the whole FRSC the next hypothesis is evaluated XBRL use enhances the audit trial information needs and increases the transparency of the FRSC. In order to analyse and verify the hypotheses the scientific methods explained further are used.

2 Domain Analysis

The second chapter analyses the reporting aspects from the reporting domain view. This chapter is subdivided in accounting process and external reporting sections and analyses the reporting aspects from the domain view which is mostly related to the area of financial accounting. Thomas [Thom1996, 1] defines the financial accounting as follows: *“Financial accounting may be defined as the process of designing and operating an information system for collecting, measur-*

ing and recording business transactions, and summarizing and communication the results of these transactions to the users to facilitate making decisions". Although only the second part of the definition is applying directly to financial reporting processes, it is not possible to understand the *FRSC* without the first part driving the creation of the financial reports. Therefore the chapter is subdivided into accounting processes and reporting processes. The accounting processes known from the accounting information systems literature are analysed and modelled, especially in regards of the input and output information and documents being transferred along the process chain. The external reporting processes are analysed and modelled having special focus on the aspects of inter-organisational production, delivery and assurance of the financial reports between various information systems. While most AIS and financial accounting literature focuses on the accounting process with the result in form of financial reports, the IFRS framework provides an overview of the users of the reports. This together with the thoroughly analysis of the German and European law concerning reporting provides enough information to model the financial reporting processes in the *FRSC*.

Accounting information flows start with the business operations generating raw report data so the business operations are the basic source of information for further reporting. The business operation processes discussed in the literature are the revenue cycle dealing with sales and cash collections, the expenditure cycle dealing with purchasing and cash disbursements, the production cycle and the human resources management and payroll cycle. The important aspect for the reporting flow is the output data produced in each of the cycles [RoSt2006, 524]. Out of different ongoing accounting cycles which consist of a number of transactions the source documents describing the transactions are produced. Examples of source documents can be cash receipts, credit card receipts, cash register tapes, customer cheque, sales invoices, supplier invoices, purchase orders, time cards, deposit slips, notes for loans, payment stubs for interest or share certificates [SuTi2004, 69-72; Thom1996, 31]. The source documents contain the monetary terms which have to be entered in the appropriate journal. The possible journals are closely related to the accounting cycles described above. Examples of the journals are sales journal, purchase journal, cash receipt, disbursement journal or the general journal [MoSt1963, 104; SuTi2004, 70; GeSH2005, 596; BaSS2005, 181]. According to Moore and Stettler [MoSt1963, 104-105] journals are characterised by the the following groups of data items entered. First group is the GL accounts where a transaction is entered later with account number, amount, posting check, and debit or credit indication. The second group contains the dates of the journal entry, names

and further explanations. The entries made in the journals are also referred to as journal vouchers [GeSH2005, 596] are posted to the ledgers. Moore and Stetter refer to the following definition for the ledger *“A book in which a summary of account is kept; the final book of record in business transaction, in which all debits and credits from the journal, etc. are placed under appropriate heads”* [MoSt1963, 114]. In most cases the ledger is referred to as the general ledger (GL) although in the practice a number of enlargements and refinements of the ledgers can be found. Apart from the GL there companies can have accounts receivable ledgers, accounts payable ledgers or subsidiary ledgers representing the corollary development of the controlling account principle. The analyses explore mainly the GL as being the base for the creation of the financial reports [MoSt1963, 113; BaSS2005, 181].

The data collected in the GL is usually coded with the use of a chart of accounts [BaSS2005, 181]. The chart of accounts specifies each type of asset, liability and owners' equity assigning a code number. Identifying the transaction and making a journal entry the code number is used to indicate the accounts affected. In the continental Europe, many countries have standardised or semi-standardised charts of accounts¹. The goal is to have standardised accounting basis used for the financial statements together with standardised presentation of financial statements. Moreover the governments can conduct statistics with a high level of internal consistency of the underlying data [SuTi2004, 66].

The transfer of the data from the journals to the ledgers is the ongoing activity. In order to create a financial report which can be distributed the entity needs to perform further processes. The trial balance created from the GL upon the chart of accounts [RoSt2006, 525] is the linkage between the ongoing accounting processes and financial report preparation processes. The trial balance *“...is the listing of all debit and credit balances in ledger accounts at the end of financial period to check that balance totals agree”* [SuTi2004, 706]. The trial balance provides a test for the accuracy of the record keeping. If one transaction will not balance then the trial balance will not balance either² [SuTi2004, 71-72]. The possible errors are corrected by means of adjusting entries to record accrued, deferred, and estimated amounts. The entries are posted to the ledger accounts and an adjusted trial balance is produced. Again, in case of errors in the trial balance, the adjusting entries are posted to the ledger accounts and another adjusted trial balance is produced.

¹ For example France and Belgium have mandatory use of the standardised chart of accounts [SuTi2004, 72].

² However it is not guaranteeing that transactions are recorded properly.

The final adjusted trial balance is the basis for the preparation of the financial statements. According to the IFRS standards the financial statement comprises balance sheet, income statement, cash flow statement, statement of changes in the equity and the notes to the financial statement [IASB2006, 692]. Already during the transfer of the closing accounts from the trial balance to the financial statement the analyses is conducted which accounts are aggregated to which financial statement disclosure. After creating the financial statement the closing entries are posted to the ledger accounts.

The last aspect of the accounting process is the distinction between the financial statement and financial report addressed often as annual report and accounts [SuTi2004, 122]. The financial statement together with the explanatory notes on the financial statement and the auditors' report on them is the first part of the financial report³. The second part is varying often in form of the discussion and the analysis of accounts. There are two other aspects of the accounting process excluded from the thesis. First is the audit of the financial report and second is the consolidation of the financial report. Both are impacting the information flow and both are considered in the analysis but not described on the detailed process level.

After the financial report is prepared the external reporting processes take place. The external reporting aspects are analysed in the further section of the chapter. The financial accounting, AIS and XBRL literature often presents simplified model of financial reporting [Hoff2006, 148; RoSt2006, 532; GeSH2005, 593; RaKO2006; HAon2005, 73; DeGr2001, 65; NuSt2002, 450]. The simplified approach requires further investigation because various report formats, various GAAPs used, various information scope in various reporting areas contradict using the model. The framework of the analysis of the external reporting is derived from the IFRS. The IFRS specifies the users of the financial statements and their information needs. According to the IFRS framework:

“The users of financial statements include present and potential investors, employees, lenders, suppliers and other trade creditors, customers, governments and their agencies and the public. They use financial statements in order to satisfy their different needs for information” [IASB2006, 35].

The government and their agencies are subdivided basing on the organisational structure known in Germany into supervising institutions, companies' registers, statistic offices and tax offices. For each of these groups an analyse is conducted based on several criteria. The criteria are di-

³ This part of the financial report is often addressed as the annual accounts.

vided into two groups. The process oriented criteria concern the reporting process itself and are the following:

- report sender - analysing details concerning the sending institution;
- reporting process - analysing the means of which a report is transmitted (postal way, fax, email, data storage medium, reporting platform) and in which format the data is being sent (paper, PDF, Word, Excel, XML, XBRL, online forms);
- reporting timeframe - focusing on the time on which the report is submitted (specified date or an initiating action);
- reporting frequency - analysing the frequency of reporting (yearly, quarterly, monthly);
- further use - analysing the use of the reports by the receiver (analysis, archiving, publication).

The second group of criteria focuses on the report as such:

- art of report - specifying the art of submitted report (single financial statement, consolidated financial statement, full financial report, quarterly financial statement, additional disclosures related to the financial statement);
- report format - analysing if the report format is predefined or freely adaptable for the sending institution;
- legal regulation - analysing the underlying legal regulation requiring the submission of the financial reports;
- auditing - analysing if the submitted report needs to be audited.

The above criteria are evaluated for each of the groups mentioned in the IFRS framework adjusted to the situation in Germany. The evaluation is based on the literature and legal requirements review as well as contact with selected institutions. The results of the evaluation and the analysis from the first chapter section are basis for the construction of the *FRSC* model in chapter 4.

3 Technical Analysis

The technical analysis of the eXtensible Business Reporting Language (XBRL) is the substantial subject of the third chapter. The discussion starts with the introduction to XBRL Specifica-

tion which is the basis documentation of the language⁴. In the first part the focus is XBRL for financial reporting (XBRL FR) which is regarded to be the core XBRL technology [Hoff2006, 16; BoWo2005, 13]. The definitions and critical analysis of terms such as XBRL taxonomies, taxonomy extensions and instance documents together with the analysis of the issues concerning the current XBRL specification [EHSK2003] build the next section of the chapter. XBRL apart from the mentioned FR adaptation has a separate adaptation for the general ledger⁵ (XBRL GL). XBRL GL plays an important role in the internal reporting domain while standardising the journal entries, GL itself and the trail balances [RaKO2006; HAon2005, 74; KrSc2003, 80] as well as in hybrid reporting discussed later in the chapter. Further and recent XBRL technologies such as XBRL Dimensional Taxonomies (XDT) allowing representation of multidimensional data sets [HRWa2006], XBRL Functions and Formulas allowing advanced calculations and validations [Hams2005, Enge2005], XBRL Versioning and XBRL Rendering are analysed and their importance for the further research is assessed. An overview and categorisation of XBRL technologies is the result of the technical analysis in the third chapter.

XBRL was first named eXtensible Financial Reporting Markup Language (XFRML) but soon the XBRL community stated that the language can have broader use and adjusted its name to incorporate various business reporting aspects [GaHa2005, 57; Hoff2006, 45; XBRL2006b]. Combining the XBRL definition from the XBRL specification and the definition of financial reporting from Wagenhofer and Ewert [WaEw2003, 4] the XBRL FR could be outlined as follows: *“XBRL for financial reporting comprises all XBRL enabled information systems oriented towards external users such as investors, creditors, customers, suppliers, competitors and public”*.

Table 1 explains the basic terms in XBRL FR area which are taxonomies and instance documents. XBRL taxonomies reflect the underlying financial reporting principles in form of different Generally Accepted Accounting Principles (GAAP)⁶ encoded using standardised XBRL vocabulary. The instance documents reflect financial statements of an entity, but in digital format.

⁴ XBRL is *de facto* standard for digital business reporting so the terms *language* and *standard* are used interchangeably [Berg2003, 15-16].

⁵ Both XBRL FR and XBRL GL are based on the XBRL Specification.

⁶ Although division between principle, rule and definition based accounting standards exists between different GAAPs the statements reflects the IFRS view as principle based accounting standard.

Table 1: Relation between Traditional Reporting and XBRL FR

	Underlying accounting principles	Financial report
Traditional reporting	GAAP	Paper, PDF or HTML ⁷ financial report
XBRL FR	GAAP based XBRL taxonomy	Instance document

The figure 1 provides a more detailed view of the XBRL FR framework. Basis terms like taxonomy, taxonomy extension, instance document or Discoverable Taxonomy Set (DTS) are visualised together with relations between them.

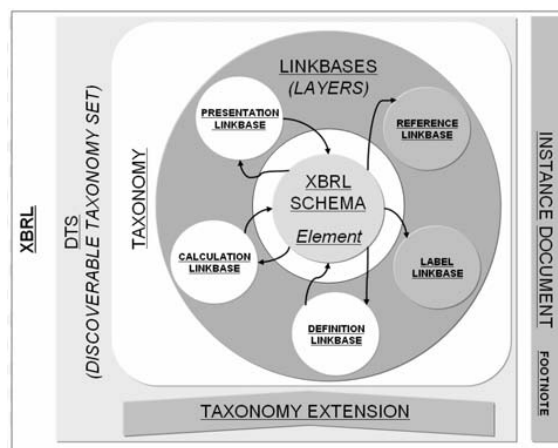


Figure 1: XBRL Financial Reporting Framework [IASC2006a]

Taxonomy in general means a catalogue or a set of rules for classification. In XBRL, taxonomy is a dictionary, containing computer-readable definitions of business reporting terms as well relationships between them and links connecting them to human-readable resources. A typical taxonomy consists of a schema (or schemas) and linkbases. A set of taxonomies that can be discovered⁸ from one entry point schema is called Discoverable Taxonomy Set (DTS) [EHSK2003, 16-17; Hoff2006, 77; IASC2006a].

⁷ Formats such as HTML and PDF are not treated as digital format even though being application consumable. HTML and PDF are not describing the reported facts in a standardised way using for example tags.

⁸ DTS discovery is a technical term and means traversing over related XBRL schemas and linkbases [EHSK2003, 16-17].

Taxonomy extensions⁹ add concepts and modify the relationships among the concepts in the base taxonomies that they extend. They are created to support specialised reporting requirements in specific accounting jurisdictions, in specific industries, or for specific companies. Taxonomy extensions consist of a set of taxonomy schemas and/or linkbases that augment a DTS that includes the base taxonomies [IASC2006a]. An instance document is a business report in XBRL format. It contains tagged business facts, together with the context in which they appear and unit description [EHSK2003, 13; IASC2006a] and is referring the tags to the elements specified in the taxonomy.

Apart from the XBRL Specification 2.1 there are more governing documents defining the rules for XBRL FR vocabulary and taxonomies architecture. The most important for creation of taxonomies is Financial Reporting Taxonomy Architecture (FRTA). FRTA states a set of 104¹⁰ rules concerning best practices of taxonomy creation [HGHH2005, 4-5]. Another set of rules which is the Financial Reporting Instance Standards (FRIS) exists for the creation of instance documents and facilitate the analysis and comparison of XBRL financial reporting data by computer applications and human readers¹¹ [GoHa2004, 1]. Finally, underlying principles for modelling of financial reporting taxonomy were created by Hoffmann [Hoff2006, 265-355]. So called patterns are a collection of 20 modelling rules which help to create standardised taxonomies which are FRTA valid.

Different taxonomies are required for different financial reporting purposes. National jurisdictions may need their own financial reporting taxonomies to reflect their local accounting regulations. Many different organisations, including regulators, specific industries or even companies, may require taxonomies to cover their own business reporting needs. The presented research considers taxonomies with the potential use in the European area with the special focus on the business reporting in Germany. The discussed taxonomies are International Financial Reporting Standards General Purpose (IFRS-GP), United States Generally Accepted Accounting Principles (US GAAP)¹² and German Accounting Principles (GermanAP).

⁹ The term taxonomy extension is used interchangeably with the term extension taxonomy [Hoff2006, 110; TeHM2003, 1-2].

¹⁰ The number of FRTA rules is changing because of the ongoing Domain Working Group (DWG) enhancements to the document.

¹¹ FRTA and FRIS similarly to XBRL Specification are accompanied by conformance suits in order to achieve greater software compatibility [Wall2004; Wall2005a].

¹² According to the assumptions of the thesis and focus on the profit-generating entities in the commerce and industry and not financial or insurance area only the US GAAP Commerce and Industry (CI) taxonomy is taken into consideration.

The XBRL GL taxonomy is intended to provide a standardized format for representing the data fields found in accounting and operation systems and transactional reports that will allow organizations to tag journal entries, accounting master files, historical status reports in XBRL and the underlying detail for financial reporting taxonomies [XBRL2005]. XBRL GL even being treated as an extra adaptation of XBRL is not a separate specification, but is based on the XBRL Specification 2.1. However, XBRL GL is not related to the FRTA and FRIS documents and conformance suits. The XII published as drafts the XBRL GL Instance Standards (GLIS) to facilitate the analysis and comparison of XBRL GL data by computer applications and human readers [Wall2005b, 3] as well as GL Taxonomy Framework Technical Architecture (GLFTA) establishing rules and conventions that assist in comprehension, usage and performance among different journal focussed taxonomies [Wall2005c, i]. From the technical point of view it is a stand-alone taxonomy, suitable for the needs of representing basic accounting databases and transactions. The most important features of the XBRL GL taxonomy according to XII are:

- multi-GAAP, drill-up to multi-XBRL reporting taxonomies capability;
- being a standard format to move unposted and posted GL information back and forth from branch offices to consolidating systems, budgeting and forecasting tools and reporting tools;
- being a standard format to move information from client systems to auditor system;
- being a tool for representing detail drill-down for performance measurement reporting items;
- creating possibilities for any type of mandatory audit trail [XBRL2006g].

The chapter 3 analyses and summarises the XBRL standard from the technical point of view. Major research is conducted using available XBRL governing documents and literature as well as a number of small prototypes. Also the practical experiences with the XBRL taxonomy development are reflected in the research results. The results of the chapter 3 are necessary for the construction of the XBRL FRSC in chapter 4.

4 XBRL Financial Reporting Supply Chain Model

Based on the definition and the analysis of both accounting processes and external reporting processes, the *FRSC* model is designed. Chapter 4 presents the *FRSC* model merging the accounting processes with the external reporting processes. The model includes the processes,

participants and information flows during production, delivery and assurance of financial reports in an inter-organisational environment. Based on the developed model, the *FRSC* definition is deduced. Parallel to the reporting discussion the auditing and assurance of financial reports is analysed [KoHs2006, 31; Sutt2006, 4]. The chapter consists also a critical consideration of the XBRL Business Reporting Supply Chain.

The methods used for the modelling approach are the event-driven process chains (EPC) known from the ARIS framework. Figures 2, 3 and 4 are the first attempt to model the accounting process and reporting process identified within the *FRSC*. In order to visualise the processes, the accounting processes are subdivided into ongoing accounting processes and report preparation processes. Ongoing accounting processes presented in the figure 2 describe the input and output documents used during repetitive accounting activities.

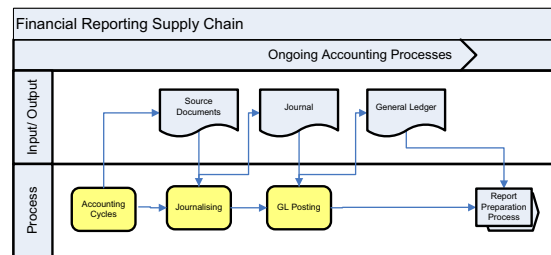


Figure 2: Ongoing Accounting Processes in the FRSC

The figure 3 presents the report preparation processes. Different than the ongoing accounting processes the report preparation is conducted usually at the end of the fiscal periods. The information included in the GL is passed through a number of processes in order to receive the audited financial report.

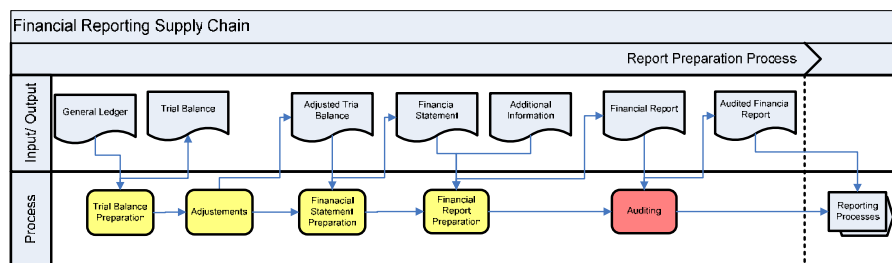


Figure 3: Report Preparation Process in the FRSC

After the financial report is prepared the reporting processes can start. The figure 4 provides a simplified model of the financial reporting processes. First important observation is that various receiving institutions require different scope of the financial report or require additional information. Also an important observation leads to a conclusion that there are receiving institutions not conducting any further processing of the received reports.

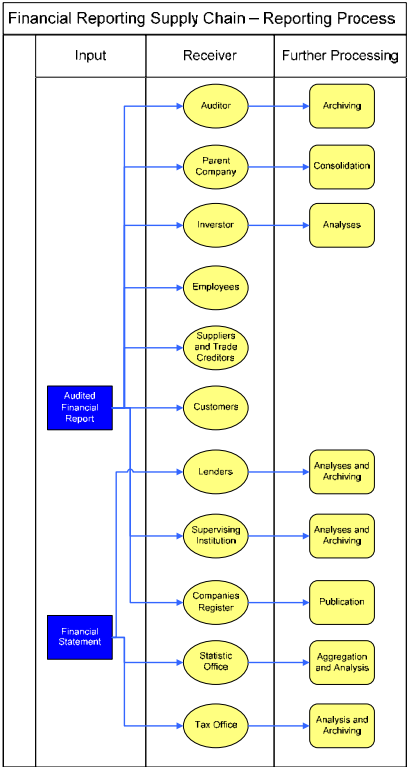


Figure 4: Reporting Processes in the FRSC

The following section of the chapter combines the three groups of processes presented above into a consistent and complete *FRSC* model. The definition of the *FRSC* is deduced from the constructed model.

The fourth chapter presents further the conceptual model of the *XBRL FRSC*. The model is based on the *FRSC* model and enhanced with the results of the technical analyses from the third chapter. The first section analyses the interaction between XBRL and different IS. The AIS, reporting systems, receiver's systems as well as inter-organisational systems are considered as fundamental parts of the *XBRL FRSC*. The *XBRL FRSC* consists of various processes which can be categorised in forward processes, concerning reporting and backward processes, concerning

both reports requests and information assurance and auditing [Sutt2006, 4]. Different participants of the *FRSC* such as management accountants, auditors, external stakeholders [BePr2003, 167-168] and their information needs are assessed and modelled in the context of XBRL technology. Chapter's result is the *XBRL FRSC* model presented in the final section.

5 Experimental Measures

In order to verify both hypotheses, not only the conceptual model of the *XBRL FRC* is constructed, but also experimental measurements are developed in the fifth chapter. The proof of concept provides results of implementing the conceptual model for a German company listed on the Frankfurt Stock Exchange. The implementation method is a prototype of the *XBRL FRC* along the whole chain from business operations to various receiving entities being internal and external stakeholders. In order to evaluate the prototype, the efficiency as well as benefits and costs analysis are taken into consideration. The prototype is generalised and its use for the whole reporting domain is analysed.

6 Conclusions

The doctoral thesis outlines the research results and conclusions in the final chapter. Therefore the dissertation evaluates the taken approach and refers if to the theses stated in the first and verified in the fifth chapter. The conclusions address also the practical use of the research results in the area of financial reporting. The consideration is going further shortly analysing the use of XBRL in various other reporting areas being out of scope of the doctoral thesis such as statistical reporting, environmental reporting, managerial reporting, tax reporting or other business reporting. The conclusions section delivers also topics where further research can be conducted.

References

- [BaSS2005] Bagranoff, N. A.; Simkin, M. G.; Strand Norman, C.: Core Concepts of Accounting Information Systems, Ninth Edition, Wiley, 2005.
- [BePr2003] Beattie, V.; Pratt, K.: Issues concerning web-based business reporting: an analysis of the views of interested parties, in: The British Accounting Review, No. 35, 2003, pp. 155-187.
- [Berg2003] Bergeron, B.: Essentials of XBRL - Financial Reporting in the 21st Century, John Wiley & Sons, New Jersey, 2003.
- [BoWo2005] Boritz, J. E.; Won, G. N.: Security in XML-based financial reporting services on the Internet, in: Journal of Accounting and Public Policy, No. 24, 2005, pp. 11-35.
- [DeGr2001] Debreceeny, R.; Gray, G. L.: The production and use of semantically rich accounting reports on the Internet: XML and XBRL, in: International Journal of Accounting Information Systems, No. 2, 2001, pp. 47-74.
- [EHSK2003] Engel, P.; Hamscher, W.; Shuetrim, G.; Kannon, D. v.; Wallis, H.: Extensible Business Reporting Language (XBRL) 2.1, XBRL International, 2003, <http://www.xbrl.org/Specification/XBRL-RECOMMENDATION-2003-12-31+Corrected-Errata-2005-11-07.rtf>, last call 18.07.2006.
- [Enge2005] Engel, P.: XBRL Functions Requirements 1.0, XBRL International, 2005, <http://www.xbrl.org/technical/requirements/Function-Req-CR-2005-06-21.rtf>, last call 09.08.2006.
- [GaHa2005] Garbellotto, G.; Hannon, N.: Why XBRL Is a “Business” Reporting Language, in: Strategic Finance, May 2005, pp. 57-61.
- [Gass2000] Gassen, J.: Datenbankgestützte Rechnungslegungspublizität, Peter Lang, Frankfurt am Main, 2000.

- [GeSH2005] Gellinas, U. J. Jr.; Sutton, S. G.; Hunton, J. E.: Accounting Information Systems, Sixth Edition, South-Western, 2005.
- [GoHa2004] Goodhand, M.; Hamscher, W: Financial Reporting Instance Standard 1.0, XBRL International, 2004, <http://xbrl.org/technical/guidance/FRIS-PWD-2004-11-14.zip>, last call 07.08.2006.
- [Hams2005] Hamscher, W: XBRL Formula Requirements, XBRL International, 2005, <http://www.xbrl.org/technical/requirements/Formula-Req-CR-2005-06-21.rtf>, last call 18.07.2006.
- [Hann2004] Hannon, N.; Why Should Management Accountants Care about XBRL?, in: Strategic Finance, July 2004, pp. 55-56.
- [HAon2005] hAonghusa, C. O.: The Journal Taxonomy, in: Accountancy Ireland, Vol. 37, No. 5, October 2005, pp. 72-74.
- [HGHH2005] Hamscher, W.; Goodhand, M.; Hoffmann, C.; Homer, B.; Macdonald, J.; Shuetrim, G.; Wallis, H.: Financial Reporting Taxonomies Architecture 1.0, XBRL International, 2005, <http://www.xbrl.org/technical/guidance/FRTA-RECOMMENDATION-2005-04-25+corrected-errata-2006-03-20.rtf>, last call 18.07.2006.
- [Hoff2006] Hoffmann, C.: Financial Reporting Using XBRL - IFRS and US GAAP Edition, Lulu, 2006.
- [HRWa2006] Hernández-Ros, I.; Wallis, H.: XBRL Dimensions 1.0, XBRL International, 2006, <http://www.xbrl.org/Specification/XDT-CR3-2006-04-26.rtf>, last call 18.07.2006.
- [IASB2006] International Accounting Standards Board: International Financial Reporting Standards 2006, IASB, London, 2006.
- [IASC2006a] International Accounting Standards Committee Foundation: Fundamentals of XBRL, 2006, http://www.iasb.org/xbrl/about_xbrl/fundamentals_xbrl.html, last call 07.08.2006.

- [KrSc2003] Kranich, P.; Schmitz, H.: Die Extensible Business Reporting Language - Standard, Taxonomien und Entwicklungsperspektiven, in: Wirtschaftsinformatik, No. 45, 2003, pp. 77-80.
- [KoHs2006] Koo, M; Hsieh, H.: Improvement of Continuous Auditing Efficiency Using Mobile and Internet Information Technologies, in: International Journal of Global Logistics & Supply Chain Management, Vol. 1, No. 1, May 2006, pp. 25-33.
- [MoSt1963] Moore, F. E.; Stettler, H.F.: Accounting Systems for Management Control, Richard D. Irwin, Homewood, 1963.
- [NuSt2002] Nutz, A.; Strauß, M.: eXtensible Business Reporting Language (XBRL) - Konzept und praktischer Einsatz, in: Wirtschaftsinformatik, No. 44, 2002, pp. 447-457.
- [RaKO2006] Ramin, K. P.; Kesselmeyer B.; Ott S.: XBRL im Internal Financial Reporting von Unternehmensgruppen, in: Zeitschrift für internationale und kapitalmarktorientierte Rechnungslegung, No. 3, March 2006, <http://www.kesselmeyer.com/blog/wp-content/uploads/2006/03/XBRL%20Internal%20Financial%20Reporting%20KRSOBK%20060301%20Lang.pdf>, last call 18.07.2006.
- [RoSt2006] Romney, B. M.; Steinbart, P. J.: Accounting Information Systems, Tenth Edition, Pearson Prentice Hall, 2006.
- [SuTi2004] Sutton, T.: Corporate Financial Accounting and Reporting, Prentice Hall, Harlow, 2004.
- [Sutt2006] Sutton, S. G.: Enterprise systems and the re-shaping of accounting systems: A call for research, in: International Journal of Accounting Information Systems, No. 7, 2006, pp. 1-6.
- [TeHM2003] Teixeira, A.; Hoffmann, C.; Macdonald, J.: Taxonomy Mapping: The Process of Creating Extension Taxonomies, 2003, http://www.iasb.org/xbrl/images/xbrllab/past_projects/CreatingExtensionTaxonomiesDraft2003-03-24.pdf, last call 08.08.2006.

- [Teix2005] Teixeira, A.: Implications of XBRL - financial reporting research opportunities, 2005, http://www.iasb.org/xbri/images/xbriiab/past_projects/xbri_paper_for_iasb_2005.pdf, last call 18.07.2006.
- [Thom1996] Thomas, A.: An Introduction to Financial Accounting, McGraw-Hill, Maidenhead, 1996.
- [WaEw2003] Wagenhofer, A.; Ewert, R.: Externe Unternehmensrechnung, Springer, 2003.
- [Wall2004] Wallis, H.: FRIS 1.0 Conformance Suit, XBRL International, 2004, <http://xbri.org/technical/guidance/FRIS-CONF-PWD-2004-11-14.zip>, last call 07.08.2006.
- [Wall2005a] Wallis, H.: FRTA 1.0 Conformance Suit, XBRL International, 2005, <http://www.xbri.org/technical/guidance/FRTA-CONF-RECOMMENDATION-2005-04-25.rtf>, last call 07.08.2006.
- [Wall2005b] Wallis, H.: XBRL GL Instance Standards 1.0, XBRL International, 2005, <http://www.xbri.org/int/gl/2005-11-07/GLIS-CR-2005-11-07.rtf>, last call 09.08.2006.
- [Wall2005c] Wallis, H.: GL Taxonomy Framework Technical Architecture 1.0, XBRL International, 2005, <http://www.xbri.org/int/gl/2005-11-07/GLTFTA-CR-2005-11-07.rtf>, last call 09.08.2006.
- [XBRL2005] XBRL International: XBRL GL 2005, 2005, <http://www.xbri.org/int/gl/2005-11-07/gl-2005-11-07.htm>, last call 09.08.2006.
- [XBRL2006b] XBRL International: XBRL's History, 2006, <http://www.xbri.org/history.aspx>, last call 07.08.2006.
- [XBRL2006g] XBRL International: XBRL GL Key Features, 2006, <http://www.xbri.org/GLKeyFeatures/>, last call 09.08.2006.

Potenzial experimenteller Forschung in der Wirtschaftsinformatik

Dipl.-Kfm. Thomas Wilde

Institut für Wirtschaftsinformatik und Neue Medien
Ludwig-Maximilians-Universität München
D-80539 München
wilde@bwl.uni-muenchen.de

Betreuender Hochschullehrer:

Prof. Dr. Thomas Hess

Institut für Wirtschaftsinformatik und Neue Medien
Ludwig-Maximilians-Universität München
D-80539 München
thess@bwl.uni-muenchen.de

Abstract

Verglichen mit der angloamerikanischen Schwesterdisziplin Information Systems ist die deutschsprachige Wirtschaftsinformatik eher konstruktionswissenschaftlich orientiert. Allerdings hat sich bisher nur ein geringes theoretisches und methodisches Fundament herausgebildet. Mit der „rigor vs. relevance“-Debatte des Information Systems Research liegt das Augenmerk seit einiger Zeit wieder stärker auf diesem Aspekt. Das nachfolgend umrissene Dissertationsvorhaben möchte Zusammenhänge zwischen Methodenprofil und theoretisch-methodischem Fundament der Disziplin untersuchen und zeigen, auf welche Weise experimentelle Forschung in der WI zur systematischen Entwicklung und Erweiterung von Theoriekernen – und damit zur kumulativen Forschung beitragen kann. Aufbauend auf eine erkenntnistheoretische Analyse gegenwärtiger Herausforderungen der Disziplin sollen Experimentaldesigns für die Wirtschaftsinformatik entwickelt und empirisch getestet werden.

1 Einführung und Problemstellung

In der Wirtschaftsinformatik (WI) und ihrer angloamerikanischen Schwesterdisziplin Information Systems (IS) finden sich grundsätzlich zwei erkenntnistheoretische Paradigmen. Das konstruktionswissenschaftliche („Design Science“) strebt nach Erkenntnisgewinn durch Schaffen und Evaluieren von IT-Lösungen in Form von Modellen, Methoden oder Systemen, die Probleme einer Organisation lösen. Das behavioristische oder verhaltenswissenschaftliche („Behavioral Science“) sieht hingegen die Analyse des Verhaltens und der Auswirkungen von existierenden Informationssystemen auf Organisationen vor. Diese beiden Paradigmen sind jeweils eng mit erkenntnistheoretischen Grundhaltungen verknüpft, die vor allem Feststellungen bezüglich möglicher Erkenntnisprozesse umfassen. Im Bereich der Wirtschaftsinformatik handelt es sich hierbei hauptsächlich um den kritischen Rationalismus und den methodischen (oder Erlanger) Konstruktivismus [Wyss04, 59-65]. Während im kritischen Rationalismus in einer objektiv gegebenen und bedingt erfassbaren Realität Theorien durch empirische Prüfung deduzierter Hypothesen systematisch auf Fehler geprüft werden (Falsifikationsversuche), ist die Aufgabe des Wissenschaftlers im methodischen Konstruktivismus anders gelagert. Hier steht die sprachliche Rekonstruktion und Analyse des Zustandekommens von menschlichen Konstruktionen im Mittelpunkt, die hier als wahrnehmbare subjektive Aspekte der Realität zu verstehen sind. Der kritische Rationalismus wird in der Wirtschaftsinformatik häufig referenziert, wenn auch zu vermuten ist, dass den hohen methodischen Anforderungen nicht immer entsprochen wird [Lehn95, 25]. Der methodische Konstruktivismus kommt in der WI vor allem in Bereichen wie der Unternehmensmodellierung oder der Anwendungssystementwicklung zum Tragen [Wyss04, 63].

Vor diesem Hintergrund führt die IS-Community seit Beginn der 90er Jahre unter dem Schlagwort „rigor versus relevance“ eine Debatte, in der die teilweise mangelnde Praxisrelevanz der verhaltenswissenschaftlich geprägten IS-Forschung kritisiert wird [BeZm03]. Als Lösungsansatz wird oft eine weniger strikte Ausrichtung auf empirisch-quantitative Methoden zugunsten von konstruktionsorientierten Vorgehensweisen propagiert [HMPR04]. So soll bei etwas geringerer Stringenz („rigor“) des Erkenntnisprozesses die Praxisrelevanz der Ergebnisse gestärkt werden.

Einerseits kommt die deutschsprachige Wirtschaftsinformatik den Vorstellungen, die im Information Systems Research bezüglich einer konstruktivistischeren Ausrichtung der Disziplin entwickelt werden bereits sehr nahe, auch gelten ihre Ergebnisse gemeinhin als praxisrelevant.

Andererseits wird der WI jedoch ein Mangel an kumulativer Forschung attestiert [Mert04, 10-12; Lang06, 21-22], der unter anderem auf die Ausrichtung an Modethemen und die geringe methodische und theoretische Fundierung in Verbindung mit mangelnder kritischer Reflexion von Forschungsergebnissen zurückgeführt wird [Goek03, 6-7; Hein01, 94-96; Lehn99, 20]. So wird durch „Prototyping auf konzeptueller Ebene“ oft ein kritisches, erkenntnistheoretisches Ziel verfehlt: die Rückkopplung mit der Theoriebasis und deren Erweiterung durch wiederholte, stringente („rigorously“) Evaluation des Forschungsergebnisses [Hein01, 94; HMPR04, 85].

Das beschriebene Dissertationsvorhaben möchte dieses Problemfeld aus einer methodologischen Perspektive bearbeiten und damit gleichzeitig den Betrachtungsschwerpunkt auf die methodische Fundierung legen. Es wird die Fragestellung bearbeitet, inwieweit die methodische Ausrichtung der Disziplin ursächlich für das Problem mangelnder theoretischer Fundierung und kumulativer Forschung ist, welche Implikationen dies nach sich zieht und welches Potenzial die Experimentalmethode als ein möglicher Berührungspunkt konstruktionsorientierter und verhaltensorientierter Forschung in diesem Kontext aufweist.

2 Konzeption

2.1 Zielsetzung und Relevanz

Eine methodologische Arbeit behandelt im Allgemeinen die Entwicklung von deskriptiven, explanatorischen oder präskriptiven Theorien oder Thesen zu den Methoden der zugrunde liegenden Disziplin. Anknüpfend an diese Dreiteilung und ergänzt um die Evaluation der so gewonnenen Ergebnisse ergeben sich für die Arbeit folgende Forschungsziele:

- **Deskription:** Charakterisierung des Problems geringer theoretischer Fundierung und kumulativer Forschung sowie des Methodenprofils der Wirtschaftsinformatik in einem adäquaten erkenntnistheoretischen Bezugsrahmen
- **Explanation:** Identifikation und Analyse von Ursachen und Mechanismen, die das beschriebene Problem betreffen, Entwicklung eines diesbezüglichen methodologischen Anforderungsprofils, Analyse von Potenzial und möglichen Problemen der Experimentalmethode im bearbeiteten Kontext

- Präskription: Diskussion und Anpassung von Varianten der Experimentalmethode im Hinblick auf das Anforderungsprofil
- Evaluation: Potenzial der analysierten Experimentaldesigns empirisch beleuchten

In Summe soll ein theoretisch fundierter, technologisch-präskriptiver Forschungsbeitrag in Form von anwendbarem Methodenwissen erarbeitet werden, der dem Anwender die Verfolgung ausgewählter wissenschaftsstrategischer Ziele erlaubt.

Die Relevanz des Forschungsvorhabens begründet sich insbesondere aus einem Spezifikum, welches durch die Schnittstellenposition der WI zwischen Wirtschafts-, Sozial- und Ingenieurwissenschaften bedingt ist. In dieser Position empfiehlt sich für die Wirtschaftsinformatik eine methodenpluralistische Ausrichtung [Fran98, 20], die sich auf internationaler Ebene z. B. in der Diskussion um die beiden dominanten erkenntnistheoretischen Paradigmen manifestiert. Die meisten Gegenstände der WI werden im Verlauf ihres Forschungsfeldlebenszyklus aus verschiedenen Perspektiven (fachlicher Fokus, Methodik, erkenntnistheoretische Position) untersucht. So ist in vielen Fällen zunächst eine informatik- bzw. implementierungsnahе Perspektive anzutreffen, die die Leistungsfähigkeit eines Forschungsobjekts für einen bestimmten Anwendungsbereich ermittelt und dieses anpasst, worauf beispielsweise eine eher organisatorisch-ökonomische Perspektive aufbauen kann. Aus dem Umstand, dass einige Untersuchungsgegenstände auf diese Weise mit der Zeit in dem Perspektivenspektrum „wandern“ und diese Perspektiven in einer Schnittstellendisziplin deutlich verschieden ausfallen, ergibt sich die hohe Bedeutung der Kohärenz von Forschungsarbeiten aus verschiedenen Perspektiven.

Da jede dieser Perspektiven mit speziellen epistemischen und methodischen Grundfragen verknüpft ist, scheint eine Bearbeitung aus methodologischer Sicht bedeutsam, da sie die gezielte Beschäftigung mit den Rahmenbedingungen erlaubt, unter denen Forschungsarbeiten verschiedener Paradigmen möglichst gut ineinander greifen können.

2.2 Forschungsumfeld

Anders als im Information Systems Research gibt es in der deutschsprachigen Wirtschaftsinformatik relativ wenige Arbeiten zu wissenschaftstheoretischen Themen. Auf den Konferenzen „Wirtschaftsinformatik“ und „Multikonferenz Wirtschaftsinformatik“ wird das Thema fast vollständig ausgeklammert [Wyss04, 53], die meisten Arbeiten finden sich in Form von Monographien, Arbeitsberichten und Herausgeberbänden (exemplarisch [Fran04],

[Holl99], [Wyss04], [BHKN03] oder [Lang06]). Das unmittelbare Forschungsumfeld des Dissertationsvorhabens ist damit zwar vergleichsweise dünn besetzt, allerdings zeigen sich oft Parallelen zum Information Systems Research. Zudem werden durch den metawissenschaftlichen Charakter weitere Gebiete berührt. Das Forschungsumfeld umschließt damit im Kern folgende drei Bereiche:

- Erkenntnistheorie der Wirtschaftsinformatik
- Methodologie der Wirtschaftsinformatik
- Experimente in soziotechnischen Systemen

Unter den ersten Bereich, „Erkenntnistheorie der Wirtschaftsinformatik“ fallen sämtliche Positionierungs- und Paradigmendiskussionen (exemplarisch [BKSW99], [BeZm03], [HMPR04]). Hieraus soll unter Einbezug allgemeiner wissenschaftstheoretischer Literatur das Analyseframework entwickelt und das Forschungsproblem eingeordnet und expliziert werden. Der zweite Bereich befasst sich mit Methoden in der Wirtschaftsinformatik – beschreibend (z. B. [KHRP96]) oder präskriptiv (z. B. [BrHW04]). Er erlaubt die Einordnung der WI in das entwickelte Framework und die theoretische Analyse von Kausalbeziehungen zwischen Methodenprofil und dem bearbeiteten Problem. Der dritte Bereich ist auf einem geringeren Abstraktionsniveau angesiedelt, da hier die zuvor generisch entwickelten Thesen angewandt und konkretisiert werden sollen. Die Experimentalmethode soll unter Rückgriff auf Grundlagen und Erfahrungen (i. S. v. Metaanalysen) zu sozial- und ingenieurwissenschaftlichen Experimenten (exemplarisch [Zimm72], [Tich98]) im entworfenen Analyserahmen eingeordnet, konkretisiert und bewertet werden.

2.3 Methodik

Erkenntnistheoretische Arbeiten weisen in der Wirtschaftsinformatik oft ein hohes Metaphysikniveau auf, mit dem sie sich de facto einer konkreten Positionsbestimmung und Kritik weitgehend entziehen [Wyss04, 53]. Im Gegensatz dazu soll das skizzierte Dissertationsvorhaben den Untersuchungsgegenstand in einem klar definierten, erkenntnistheoretischen Bezugsrahmen bearbeiten. Insgesamt wird eine kritisch-rationalistische Grundhaltung eingenommen. Demnach sind zunächst deskriptive, explanatorische und präskriptive Thesen zu entwickeln, die anschließend, zumindest bezüglich der relevanten

Punkte, empirisch überprüft werden. Das Ergebnis besteht in dem Teil der entwickelten Thesen, die empirisch nicht widerlegt werden konnten, also als „vorläufig bewährt“ gelten.

Zur weiteren Differenzierung können drei methodische Teile der Arbeit unterschieden werden: Die Analyse der Problemstellung und die Entwicklung von Lösungsszenarien erfolgen theoretisch-deduktiv, wobei einzelne Aspekte, wie die Beschreibung des WI-Methodenprofils durch empirisch-explorative Literaturanalysen gestützt werden. Der präskriptive Teil soll sich zudem verstärkt auf Analogien mit anderen experimentierenden Disziplinen stützen. Der Evaluationsteil hat den Charakter eines „Meta-Experiments“: Hier ist vorgesehen, in einem ex-post-facto Design bereits durchgeführte Analysen zu einem Untersuchungsgegenstand inhaltsanalytisch auszuwerten und zudem im Feld verschiedene Forschungsdesigns auf denselben Gegenstand anzuwenden. Stimulus ist die Auswahl einer speziellen methodischen Vorgehensweise, Response die Ausprägung der designbezogenen Qualitätskriterien des Forschungsergebnisses. Der Vergleich der Ergebnisqualität ist in so einem Fall anhand qualitativer Kriterien, aber z. B. auch quantitativ, anhand von Varianz- und Kovarianzanalysen möglich [KuTG94].

Die Feldstudien sollen in Kooperation mit der Home Shopping Europe GmbH & Co KG, die den Teleshoppingsender HSE24 betreibt, an dem Themenfeld „interactiveTV im Teleshopping“ durchgeführt werden. Das Evaluationsszenario wird im Folgenden kurz vorgestellt.

2.4 Evaluationsszenario „interactiveTV im Teleshopping“

Bei Teleshoppingsendern handelt es sich prinzipiell um Versandhändler, die mit ihren Kunden über mehrere Kanäle, vorwiegend via TV und Callcenter, aber auch via Internetshops und SMS in Verbindung stehen. In der Praxis stellt sich vor dem Hintergrund des Ausbaus der Breitbandnetze und Innovationen wie IPTV die Frage, ob ein weiterer Kommunikationskanal „interaktives Fernsehen“ (iTV) eingeführt werden sollte. Forschungsziele sind hier die Evaluation verschiedener Varianten von iTV-Shopping (technische Möglichkeiten, Nutzerakzeptanz, Kaufverhalten, etc.) sowie die Analyse von ökonomischen Implikationen für das gesamte Unternehmen. Beispielsweise ist mit medienübergreifenden Änderungen im Kundenverhalten, insbesondere mit Kannibalisierungseffekten zu rechnen.

Es ist geplant, eine veränderbare Prototypapplikation zum iTV-Shopping zu entwickeln, die als Kern verschiedener Untersuchungsdesigns zur Klärung der fachlichen Fragen, aber auch zur Evaluation der im Dissertationsvorhaben zu entwickelnden Thesen dienen soll. Die Analyse gliedert sich in drei Phasen:

- Phase I: Entwicklung und Bewertung von Varianten des iTV-Teleshopping
➔ Durchführen einer theoretischen Analyse verschiedener Varianten des iTV-Shopping und Abbildung dieser in einem veränderbaren Prototyp.
- Phase II: Validierung der bisherigen Ergebnisse und Analyse der Implikationen
➔ Analyse von Kausalzusammenhängen zwischen Applikationsspezifika (z. B. Warenkorbfunktion, 1-click-buy, Empfehlungssysteme, etc.) und Kundenverhalten im Experiment.
- Phase III: Metaanalyse des Forschungsdesigns

Im Rahmen der Metaanalyse sollen für jede Phase einzeln und phasenübergreifend Qualitätskriterien (interne und externe Validität, Reliabilität, Objektivität) diskutiert und den im Dissertationsvorhaben zuvor entwickelten Thesen gegenübergestellt werden.

3 Lösungsansätze und erste Überlegungen

3.1 Überblick

Anknüpfend an die Zielsetzung ist das Vorhaben in vier inhaltliche Komponenten gegliedert: Auswahl des Analyseframeworks und Problembeschreibung, Problemanalyse, Entwicklung von Experimentaldesigns und Thesen zu deren Potenzial sowie die Evaluation dieser Ergebnisse. Die Problemanalyse soll in dem zuvor ermittelten Analyseframework Ursachen und Mechanismen identifizieren und daraus ein methodologisches Anforderungsprofil entwickeln. Das Experiment wird nun als Lösungsansatz anhand des Anforderungsprofils bewertet. Darauf aufbauend, und im Vergleich mit anderen experimentierenden Disziplinen sollen geeignete Designs für die Wirtschaftsinformatik entwickelt werden. Dies umfasst die Untersuchungsanordnung, die Auswahl von konstruktiven Techniken zur Erstellung des Evaluationsgegenstands (z. B. spezielle Varianten des Prototyping) und die Auswahl von verhaltenswissenschaftlichen Techniken (z. B. Befragung, Beobachtung) zur experimentellen Evaluation. Die Ergebnisse werden zu Thesen verdichtet und im letzten Schritt empirisch untersucht. Abbildung 1 fasst diese Vorgehensweise schematisch zusammen.

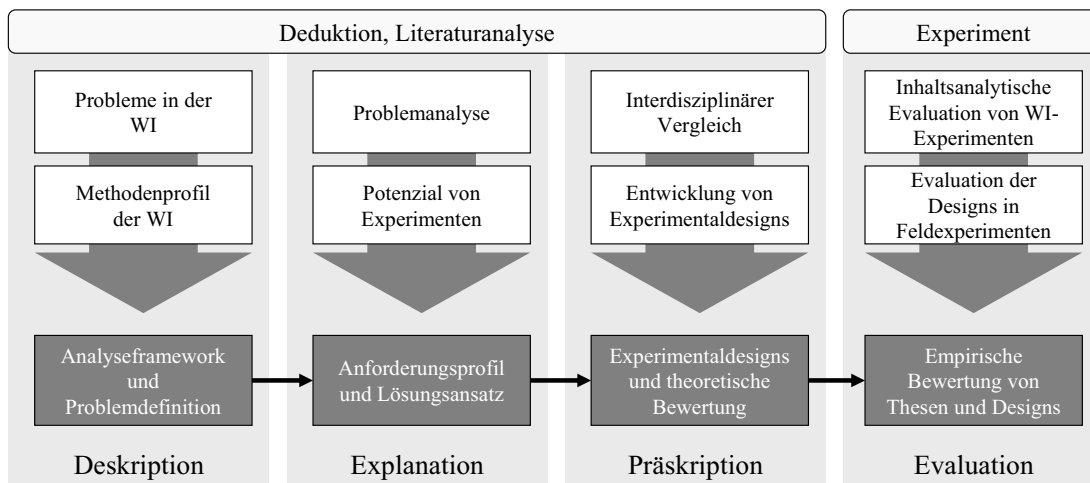


Abb. 1: Inhaltlicher und methodischer Aufbau des Dissertationsvorhabens

3.2 Begriffsverständnis Methode

Da in den folgenden Abschnitten die Repräsentation, Modifikation und Evaluation von Forschungsmethoden behandelt wird, soll zunächst das zugrunde gelegte Begriffsverständnis expliziert werden.

Unter einer Methode wird allgemein eine Art und Weise des Vorgehens verstanden, die sich durch eine bestimmte Auswahl von Instrumenten als Mittel der Zielerreichung auszeichnet [Chmi94, 36-37]. Ist ein solches Vorgehen durch intersubjektiv nachvollziehbare und nachprüfbar Verhaltensregeln beschrieben, kann grundsätzlich von einer wissenschaftlichen Methode gesprochen werden. Bisher gibt es keine disziplinübergreifende, einheitliche Definition von Forschungsmethoden, welche im Sinne von konstituierenden Merkmalen oder Taxonomien über das skizzierte, grundlegende Verständnis hinausgeht. Für die Sozialwissenschaften findet sich bei [Herr99, 36-37] eine mögliche Abgrenzung, die auch für die Wirtschaftsinformatik geeignet scheint. Er definiert wissenschaftliche Methoden als

- mitteilbare Systeme von Regeln
- die von Akteuren als Handlungspläne zielgerichtet verwendet werden können,
- intersubjektive Festlegungen zum Verständnis der Regeln und der darin verwendeten Begriffe enthalten und

- deren Befolgung oder Nichtbefolgung aufgrund des normativen und präskriptiven Charakters der Regeln feststellbar ist.

Aus der systematischen Verknüpfung der Regeln folgt, dass Beginn und Ende der Anwendung einer Methode definiert sind und verschiedene Methoden in einem hierarchischem Verhältnis zueinander stehen können.

3.3 Analyseframework und Spezifika der Wirtschaftsinformatik

Als vergleichsweise junge, im Sinne Kuhns vor-paradigmatische Wissenschaft [Wyss04, 54], bietet die Wirtschaftsinformatik kaum Anhaltspunkte für einen eigenen Bezugsrahmen zur Bearbeitung von methodologischen Fragen. Aus diesem Grund ist im vorliegenden Kontext auf ein metawissenschaftliches Modell zurückzugreifen, welches die Spezifika des betrachteten Problemfelds und des WI-Methodenprofils abbilden und verknüpfen kann. In einem ähnlichen Kontext schlägt Holl verschiedene wissenschaftstheoretische Theoriekerne, wie z. B. die evolutionäre Erkenntnistheorie oder Poppers Drei-Welten-Modell vor [Holl99, 96].

Die bereits einführend genannten, potenziellen Problemursachen sollen in diesem Bezugsrahmen untersucht werden, um Mechanismen zu identifizieren, die sie mit dem Problemkern verknüpfen. Darauf aufbauend sind Wirkungsketten zu untersuchen, die typische Eigenschaften von WI-Forschungsmethoden und -Theorien mit dem entwickelten Konstrukt verknüpfen. Diese Wirkungsketten stellen den Anknüpfungspunkt für die spätere Betrachtung des Potenzials von Experimenten dar. Die Struktur dieses Abschnitts ist schematisch in Abbildung 2 dargestellt:

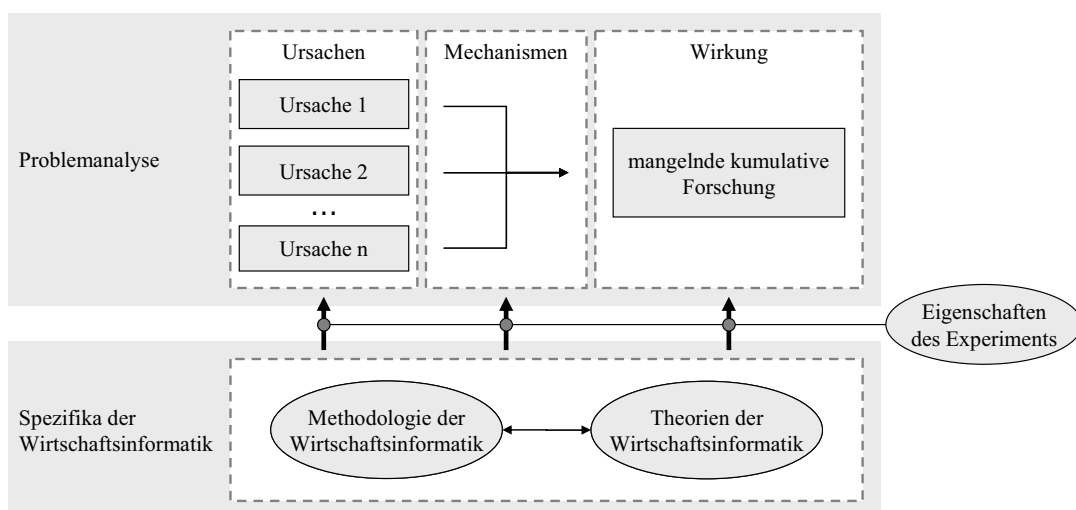


Abb. 2: Schematische Darstellung der Problemstruktur

Im Folgenden werden die Spezifika der Wirtschaftsinformatik in diesem Kontext behandelt.

3.3.1 Methodenprofil der Wirtschaftsinformatik

Die Analyse des Methodenprofils der Wirtschaftsinformatik ist ein Baustein der Problemanalyse und zum Teil bereits abgeschlossen. Einen wichtigen Eckpfeiler des Selbstverständnisses der Wirtschaftsinformatik stellt eine methodenpluralistische Erkenntnisstrategie dar, die sich Instrumenten aus den Real-, Formal- und Ingenieurwissenschaften bedient [WKWI94; Fran98, 20]. Es existieren wenige Arbeiten zu diesem Thema, weshalb eine literaturbasierte Analyse durchgeführt wurde, die die einzelnen Methoden voneinander abgegrenzt, auf ein einheitliches Abstraktionsniveau bringt und nach erkenntnistheoretischen Kriterien in ein Portfolio einordnet [WiHe06].

Hierbei wurde aus Lehrbüchern und Methodenbeiträgen eine Liste von Forschungsmethoden zusammengestellt, die in der Wirtschaftsinformatik typischerweise zum Einsatz kommen. Um Vergleichbarkeit herzustellen wurden Methoden, die nicht allen Definitionskriterien (vgl. 3.2) genügen, eliminiert und ggf. Methoden zusammengefasst oder aufgespaltet. Weiterhin wird eine Methode nur dann als eigenständig aufgeführt, wenn anhand des Regelsystems ein abgeschlossener Forschungsbeitrag erarbeitet werden kann [WiHe06, 6-8].

Die Wahl von Beschreibungsdimensionen für das Methodenportfolio ist zielabhängig und unterliegt technisch der Restriktion, dass die Dimensionen zumindest ordinal skaliert und hinsichtlich der Methodenpositionierung eindeutig sein müssen. Hierbei bestehen mehrere Möglichkeiten. Für den vorliegenden Fall wurden der Formalisierungsgrad und das erkenntnistheoretische Paradigma als Beschreibungsdimensionen gewählt, um den modalen Charakter des Methodenprofils herauszuarbeiten. Die Einordnung der Methoden wurde unter Rückgriff auf die einschlägige Methodenliteratur durch Rangfolgebildung vorgenommen. Zur Diskussion der Einordnung vgl. [WiHe06, 10-12].

Abbildung 3 zeigt das entwickelte Portfolio, welches im nächsten Schritt stichprobenartig mit empirischen Daten unterfüttert werden soll, um die methodischen Schwerpunkte und deren Entwicklung im Zeitverlauf zu identifizieren.

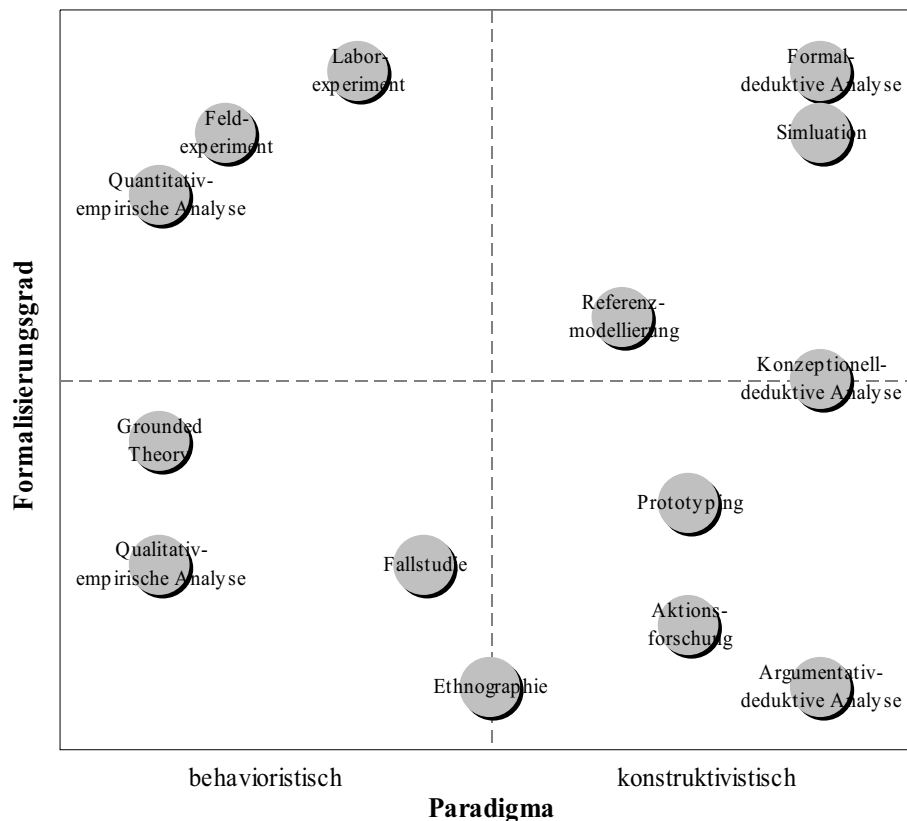


Abb. 3: Methodenspektrum der Wirtschaftsinformatik

3.3.2 Spezifika von Theorien in der Wirtschaftsinformatik

Hinsichtlich ihrer Zielsetzung können fünf Arten von Theorien in der Wirtschaftsinformatik (respektive Information Systems Research) unterschieden werden [Greg06, 614]: Analysetheorien, Erklärungstheorien, Prognosetheorien, kombinierte Erklärungs-/Prognosetheorien und Gestaltungstheorien. Sie weisen generell eine bestimmte Repräsentationsform auf und setzen innerhalb eines bestimmten Geltungsbereichs Konstrukte zueinander in Verhältnis [Greg06, 620]. Hierbei entwickelt die WI – wie Sozial- und Wirtschaftswissenschaften auch – keine eigenen Gesetzmäßigkeiten als theoretisches Fundament [Gade97, 13-17]. Sie versucht, basierend auf physikalisch-technischen Tatsachen und Gesetzmäßigkeiten anderer Disziplinen unvollständige Theorien im Sinne von ceteris-paribus-Hypothesen aufzustellen. So ist bspw. nicht davon auszugehen, dass die Einführung eines neuen ERP-Systems gesetzesartig immer identische Implikationen nach sich zieht, vielmehr ist das Geflecht an Nebenbedingungen und Kausalketten interessant, die in Summe das Theoriekonstrukt ausmachen. Dieses Charakter-

istikum ist mit zwei Problemfeldern verknüpft: Zum einen hängt der Geltungsbereich und damit auch der Nutzen einer Theorie von der zeitlichen und räumlichen Stabilität dieses Nebenbedingungsgeflechts ab. Zum anderen wird durch diese Struktur eine empirische Prüfung erheblich komplizierter, da der Kausalzusammenhang durch die unvollständige Spezifikation der Rahmenbedingung nur schwer greifbar ist.

3.4 Potenzial von Experimenten in der Wirtschaftsinformatik

Die Sozial- und Wirtschaftswissenschaften, aber auch die Informatik beantworten die Frage nach der empirischen Überprüfbarkeit unvollständiger Theorien generisch mit dem Verweis auf die Experimentalmethode und ihre Vielzahl an Kontrolltechniken [Gade97, 18; Tich98, 40; Zimm72, 66-75]. Das Experiment ist ein Design zur empirischen Untersuchung von Kausalzusammenhängen, bei dem in kontrollierter Umgebung eine Experimentalvariable auf wiederholbare Weise manipuliert (Stimulus) und die Wirkung der Manipulation (Response) gemessen wird [Zimm72, 37]. Die Umgebungskontrolle zwischen der Ausgangssituation vor der Manipulation und der Endsituation soll für die Untersuchung ein möglichst vollkommenes *ceteris-paribus* Verhältnis sicherstellen, um die gemessene Wirkung auf den Stimulus kausal zurückführen zu können (vgl. Abbildung 4).

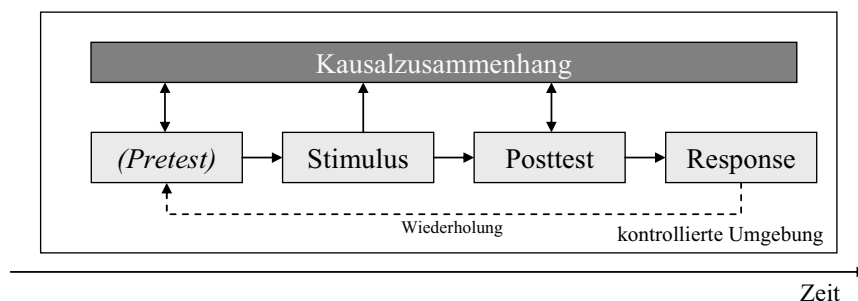


Abb. 4: Schematische Darstellung der Experimentalmethode

Damit ist die Experimentalmethode in der Lage, das erkenntnistheoretische Problem der empirischen Evaluation von WI-Forschungsergebnissen teilweise zu lösen: zwar sind die Rahmenbedingungen hier nicht vollständig spezifiziert, sie werden aber durch die Umgebungskontrolle größtmöglich konstant gehalten. Ein typisches Beispiel aus der Wirtschaftsinformatik ist ein Experiment von Jung et al., die Probanden leicht unterschiedliche Reengineeringaufgaben (Stimulus) lösen ließen, um aus deren Verhalten Rückschlüsse auf Kosteneinflussfaktoren zu ziehen [JRST96].

Der Blick auf andere Disziplinen zeigt, dass das Experiment ähnliche Rollen einnimmt. Auch wenn diese Betrachtung noch erheblich zu vertiefen ist, kann festgehalten werden, dass Experimente z.B. in Organisation [Pico72, 256-257], Kapitalmarktforschung [LiBN02] oder Marketing [KuTG94] in teilweise vergleichbaren Situationen bereits einen hohen Stellenwert innehaben. Die Wirtschaftsinformatik hat an dieser Stelle gegenüber anderen Disziplinen sogar einen Vorteil, da viele Untersuchungen das Verhalten von soziotechnischen Systemen betreffen [Fran98, 93; KHRP96]. Da die Experimentalumgebung somit häufig ein IT-System umfasst, kann der Forscher dieses als „verlängerten Arm“ nutzen. Es stellt aus dessen Sicht einen Multiplikator für die Eingriffs- und Kontrollmöglichkeiten dar und macht die physische Präsenz des Untersuchungsleiters überflüssig.

Hinsichtlich des Potenzials soll nochmals auf die Bedeutung der Verknüpfung von konstruktions- und verhaltenswissenschaftlicher Perspektive in der WI eingegangen werden. Diese Verknüpfung ist dann wichtig, wenn starke Interdependenzen zwischen Artefaktkonstruktion und dem Verhalten des soziotechnischen Systems bestehen. Dies ist z. B. dann der Fall, wenn eine der Perspektiven Nebenbedingungen berücksichtigen muss, die erst durch die andere Perspektive entwickelt werden können. Beispiele hierfür sind die weit reichenden organisatorischen Implikationen der Einführung von Standardsoftware in Unternehmen oder marktliche Implikationen der Informationssystemgestaltung im B2C-Bereich.

In einem idealisierten Szenario kann die Theorie vollständig die Schnittstellenfunktion zwischen verhaltens- und konstruktionswissenschaftlicher Perspektive erfüllen, sodass ein Forschungs- bzw. Lernzyklus zwischen beiden Perspektiven entsteht (vgl. Abbildung 5).

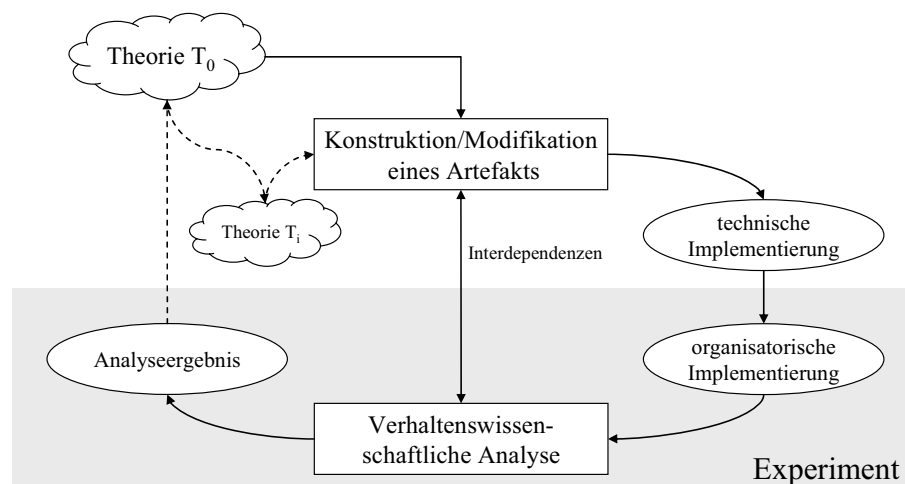


Abb. 5: Berührungspunkt von konstruktions- und verhaltenswissenschaftlicher Forschung

Diese Überlegung soll anhand des bereits angesprochenen Fallbeispiels iTV-Teleshopping illustriert werden: Anhand von Theorien zu Technologieakzeptanz und zu den im Teleshopping typischen Impulskaufentscheidungen wird ein iTV-System modelliert und prototypisch realisiert. Die experimentelle Evaluation zeigt nun, dass sich nicht das gewünschte Konsumentenverhalten einstellt. Das System wird zwar genutzt, die bereit gestellte Warenkorbfunktion stört allerdings den Impulskaufeffekt und ein hoher Anteil der Bestellvorgänge wird abgebrochen. Eine theoretische Annahme zum Impulskauf kann nun korrigiert und der Prototyp entsprechend angepasst werden. Über das Theoriekonstrukt findet eine experimentell gelenkte Evolution des Systemmodells statt.

Im explanatorischen und präskriptiven Teil der Dissertation soll der Analyseschwerpunkt auf folgenden Punkten liegen:

- Mögliche Strukturen von Theorien als Perspektivenschnittstelle
- Methoden der theoriegeleiteten Artefaktentwicklung und Implementierung
- Methoden der organisatorischen Implementierung
- Messverfahren zur experimentellen Analyse soziotechnischer Systeme
- Methoden zur Analyse und Theorierückkopplung
- Analyse der Interdependenzen zwischen diesen Problemfeldern

4 Zusammenfassung

Das beschriebene Dissertationsvorhaben möchte Zusammenhänge zwischen Methodenprofil und theoretisch-methodischem Fundament der Disziplin untersuchen und zeigen, auf welche Weise experimentelle Forschung in der Wirtschaftsinformatik zur systematischen Entwicklung und Erweiterung von Theoriekernen und damit zu kumulativer Forschung beitragen kann. Aufbauend auf die erkenntnistheoretische Analyse gegenwärtiger Herausforderungen der Disziplin sollen Experimentaldesigns für die Wirtschaftsinformatik entwickelt, deren Potenzial mit Thesen konturiert und in einem Evaluationsschritt empirisch geprüft werden. Die Arbeit soll einen erkenntnistheoretisch fundierten, technologisch-präskriptiven Forschungsbeitrag in Form von anwendbarem Methodenwissen erarbeiten, der die Verfolgung ausgewählter wissenschaftsstrategischer Ziele erlaubt.

Literaturverzeichnis

- [BeZm03] Benbasat, Izak; Zmud, Robert W.: The Identity Crisis within the IS Discipline: Defining and Communicating the Discipline's Core Properties. In: MIS Quarterly 27. Jg. (2003) Nr. 2, S. 183-194.
- [BHKN03] Becker, Jörg; Holten, Roland; Knackstedt, Ralf; Niehaves, Björn: Forschungsmethodische Positionierung in der Wirtschaftsinformatik - epistemologische, ontologische und linguistische Leitfragen. Institut für Wirtschaftsinformatik, Westfälische Wilhelms-Universität Münster, Münster 2003.
- [BKSW99] Becker, Jörg; König, Wolfgang; Schütte, Reinhard; Wendt, Oliver; Zelewski, Stephan (Hrsg.): Wirtschaftsinformatik und Wissenschaftstheorie: Bestandsaufnahme und Perspektiven. Gabler, Wiesbaden 1999.
- [Bich06] Bichler, Martin: Für Sie gelesen: Design Science in Information Systems Research, by A. Hevner, S. T. March, J. Park, S. Ram. In: Wirtschaftsinformatik 48. Jg. (2006) Nr. 2, S. 133-142.
- [BrHW04] Braun, Christian; Hafner, Martin; Wortmann, Felix: Methodenkonstruktion als wissenschaftlicher Erkenntnisansatz, Arbeitsbericht des Instituts für Wirtschaftsinformatik. Universität St. Gallen, St. Gallen 2004.
- [Chmi94] Chmielewicz, Klaus: Forschungskonzeptionen der Wirtschaftswissenschaft. 3. Auflage, Schäffer-Pöschel, Stuttgart 1994.
- [Fran98] Frank, Ulrich: Wissenschaftstheoretische Herausforderungen der Wirtschaftsinformatik. In: Gerum, Elmar (Hrsg.): Innovation in der Betriebswirtschaftslehre. Gabler, Wiesbaden 1998, S. 91-118.
- [Fran04] Frank, Ulrich: Wissenschaftstheorie in Ökonomie und Wirtschaftsinformatik. Dt. Univ.-Verl., Wiesbaden 2004.
- [Greg06] Gregor, Shirley: The Nature of Theory in Information Systems. In: MIS Quarterly 30. Jg. (2006) Nr. 3, S. 611-642.

- [Gade97] Gadenne, Volker: Erfahrung, Erkenntnis und Wirklichkeitsgestaltung - Anmerkungen zur Rolle der Empirie in der Wirtschaftsinformatik. In: Grün, O. and Heinrich, Lutz J. (Hrsg.): Wirtschaftsinformatik: Ergebnisse empirischer Forschung. Wien, u.a. 1997, S. 7-20.
- [Goek03] Goecken, Matthias: Die Wirtschaftsinformatik als anwendungsorientierte Wissenschaft. Symptome, Diagnose und Therapievorschlge. Institut fr Wirtschaftsinformatik, Phillips-Universitt Marburg, Marburg 2003.
- [Hein01] Heinrich, Lutz J.: Wirtschaftsinformatik - Einfhrung und Grundlegung, 2. Auflage, Mnchen u.a. 2001.
- [Herr99] Herrmann, Theo: Methoden als Problemlsungsmittel. In: Roth, Erwin and Holling, Heinz (Hrsg.): Sozialwissenschaftliche Methoden: Lehr- und Handbuch fr Forschung und Praxis. Oldenbourg 1999, S. 20-48.
- [HMPR04] Hevner, Alan R.; March, Salvatore T.; Park, Jinsoo; Ram, Sudha: Design Science in Information Systems Research. In: MIS Quarterly 28 (2004) 1, S. 75-105.
- [Holl99] Holl, Alfred: Objektorientierung und Poppers Drei-Welten-Modell als Theoriekerne in der Wirtschaftsinformatik. In: Schtte, Reinhard; Siedentopf, Jukka; Zelewski, Stephan (Hrsg.): Wirtschaftsinformatik und Wissenschaftstheorie - Grundpositionen und Theoriekerne, Arbeitsbericht des Institut fr Produktion und Industrielles Informationsmanagement, Universitt Essen 1999, S. 91-105.
- [JRST96] Jung, Reinhard; Rimpler, Andreas; Schnieder, Thomas; Teubner, Alexander: Ein Experiment zur Untersuchung ausgewhlter Kosteneinflufaktoren bei Reengineering-Projekten. In: Wirtschaftsinformatik 38. Jg. (1996) Nr. 2, S. 181-190.
- [KHRP96] Knig, W.; Heinzl, A.; Rumpf, M.; Von Poblitzki, A.: Zur Entwicklung der Forschungsmethoden und Theoriekerne der Wirtschaftsinformatik in den nchsten zehn Jahren. Eine kombinierte Delphi- und AHP-Untersuchung. In:

- Heilmann, Heidi (Hrsg.): Information Engineering. Oldenbourg, München u.a. 1996, S. 35-66.
- [KuTG94] Kuhfeld, W. F.; Tobias, Randall D.; Garratt, Mark: Efficient experimental design with marketing research applications. In: Journal of Marketing Research (JMR) 31. Jg. (1994) Nr. 4, S. 545-557.
- [Lang06] Lange, C.: Entwicklung und Stand der Disziplinen WI und IS. ICB Research Report Nr. 4, Institut für Informatik und Wirtschaftsinformatik, Universität Duisburg-Essen, Essen 2006.
- [Lehn05] Lehner, F.: Grundfragen und Positionierung der Wirtschaftsinformatik. In: Lehner, F., Hildebrand, K., Maier, R. (Hrsg.): Wirtschaftsinformatik – Theoretische Grundlagen. Haser, München, Wien, S. 1-72.
- [Lehn99] Lehner, F: Theoriebildung in der Wirtschaftsinformatik. In: Becker, Jörg, König, Wolfgang, Schütte, Reinhard, Wendt, Oliver and Zelewski, Stephan (Hrsg.): Wirtschaftsinformatik und Wissenschaftstheorie: Bestandsaufnahme und Perspektiven. Wiesbaden 1999, S. 5-21.
- [LiBN02] Libby, Robert; Bloomfield, Robert; Nelson, Mark W.: Experimental research in financial accounting. In: Accounting, Organizations & Society 27. Jg. (2002) Nr. 8, S. 775-810.
- [Mert04] Mertens, Peter: Zufriedenheit ist die Feindin des Fortschritts - ein Blick auf das Fach Wirtschaftsinformatik. Arbeitspapier Nr. 4/2004 des Bereichs Wirtschaftsinformatik I der Universität Erlangen-Nürnberg.
- [Pico72] Picot, Arnold: Grundfragen experimenteller Organisationsforschung: Ein wissenschafts- und methodentheoretischer Beitrag zur empirischen Betriebswirtschaftslehre. 1972.
- [Tich98] Tichy, Walter F.: Should computer scientists experiment more? In: IEEE Computer 31. Jg. (1998) Nr. 8, S. 32-40.

- [WiHe06] Wilde, T.; Hess, T.: Methodenprofil der Wirtschaftsinformatik. Arbeitsbericht 2/2006 des Instituts für Wirtschaftsinformatik und Neue Medien der Ludwig-Maximilians-Universität München.
- [WKWI94] WKWI: Profil der Wirtschaftsinformatik. Ausführungen der Wissenschaftlichen Kommission der Wirtschaftsinformatik. In: Wirtschaftsinformatik 36. Jg. (1994) Nr. 1, S. 80-81.
- [Wyss04] Wyssusek, Boris: Methodologische Aspekte der Organisationsmodellierung in der Wirtschaftsinformatik - ein soziopragmatisch-konstruktivistischer Ansatz, Dissertation an der Fakultät für Elektrotechnik und Informatik, TU Berlin 2004.
- [Zimm72] Zimmermann, Ekkart: Das Experiment in den Sozialwissenschaften. Stuttgart 1972.

Optimale Preisgestaltung von internetbasierten Diensten

Dipl. Wirt-Inf. Christian Schlereth
Professur für Electronic Commerce
Universität Frankfurt
60325 Frankfurt
schlereth@wiwi.uni-frankfurt.de

Betreuender Professor
Prof. Dr. Bernd Skiera
Professur für Electronic Commerce
Universität Frankfurt
60325 Frankfurt
skiera@wiwi.uni-frankfurt.de

Eingereicht für das Doctorial Consortium 26.-27. Februar 2007, Karlsruhe

Abstract

Jüngste technische Fortschritte im Bereich Software Engineering haben bewirkt, dass verstärkt internetbasierte Dienste in Geschäftsanwendungen oder Betriebsabläufen kostenpflichtig zum Einsatz kommen. Für die Anbieter dieser Dienste eignet sich besonders die Preisbildung in Abhängigkeit der nachgefragten Menge als preispolitisches Instrument. Unklarheit besteht jedoch über die Bestimmung der gewinnoptimalen Preishöhe, Struktur und das erwartete Verhalten der Konsumenten in solchen Tarifen. Ein in der Betriebswirtschaft bewährtes Verfahren ist die Erhebung individueller Zahlungsbereitschaftsfunktionen und das Ableiten des Nutzungsverhaltens sowie der Nachfragemenge von Kunden. Letztendlich gestatten diese Informationen die Bestimmung von gewinnoptimalen, optionalen Tarifen. Ziel dieser Dissertation aus Sicht der Wirtschaftsinformatik bildet die Fragestellung, wie die zugehörigen Modelle und Vorgehensweisen kosten-/nutzenoptimal von der technischen Seite in bestehende Infrastrukturen (z. B. Internetauftritt eines Serviceproviders) integriert werden können.

1 Ausgangssituation

Jüngste technische Fortschritte und Trends im Bereich Software Engineering haben bewirkt, dass verstärkt internetbasierte Dienste (z.B. Web Services oder Dienste zur Bereitstellung von Informationsgütern) plattformunabhängig in Geschäftsanwendungen oder in Betriebsabläufen zum Einsatz kommen. Hieraus bieten sich für Unternehmen vielfältige, neuartige Möglichkeiten zur Verwirklichung von wirtschaftlich selbst tragenden Geschäftsmodellen, in denen die Dienste kommerziell betrieben und mit den Kunden abgerechnet werden.

Diese Dienste werden üblicherweise nicht einmalig, sondern wiederkehrend von den Kunden in Anspruch genommen. Von der technischen Seite sind die Anbieter dieser Dienste mit einfachen Mitteln in der Lage, individuelle Nutzungsverhalten in Form bestimmter Leistungsmerkmale – beispielsweise Rechenleistung, Speicherkapazität oder Anzahl der Transaktionen – zu erfassen (vgl. [TaCh05]). Dadurch eignen sich für kommerzielle Dienste besonders transaktionsbasierte Erlösmodelle, welche die Höhe des Rechnungsbetrags in Abhängigkeit des tatsächlichen Nutzungsverhaltens berechnen [MaVa95]. Neben der Messbarkeit des Nutzungsverhaltens begünstigen ferner die intrinsischen Eigenschaften „Nichtlagerbarkeit“ und „Nichtübertragbarkeit“ die mengenbezogene Preisdifferenzierung (auch Preisdifferenzierung zweiten Grades oder nichtlineare Preissetzung genannt) als preispolitisches Instrument [Skie99]. Hierbei werden unterschiedliche Preise pro Mengeneinheit abgerechnet, die in Abhängigkeit der nachgefragten Mengen variieren.

In den letzten Jahren wurde die Anwendung von mengenbezogener Preisdifferenzierung bei internetbasierten Diensten in der Praxis immer populärer [Dana02]. Beispielsweise kann das Routenplanungssystem Microsoft Map Point neben dem kompletten Erwerb als Softwareanwendung auch im Sinne einer Mietsoftware über WebServices in eigene Geschäftsanwendungen eingebunden werden. Ein entsprechender Anbieter ist epidirect.com (siehe <http://www.epidirect.com/MS/MapPoint/SFees.htm>). Die Nutzung seiner WebServices ist nicht kostenlos, sondern wird über verschiedene frei wählbare Preismodelle abgerechnet: Zu diesen Preismodellen gehören ein dreiteiliger Tarif, bestehend aus monatlicher Grundgebühr, erlaubtes Nutzungskontingent und zusätzliche Preise bei Überschreitung des Nutzungskontingents, ein Wasserfallmodell, bestehend aus einer Mindestnutzungsverpflichtung sowie ein Abonnementmodell, bestehend aus einem Jahresbeitrag und einem maximalen Nutzungskontingent. Lediglich der Zugriff auf den Testserver ist zu Entwicklungszwecken kostenlos. Innerhalb der einzelnen Modelle existieren weitere auf das Verwendungsprofil des Nutzers (z.B. Gelegen-

heitsnutzer oder Vielnutzer) abgestimmte Tarife. Ein potenzieller Kunde muss sich zunächst für eins der drei Preismodelle entscheiden und anschließend denjenigen Tarif wählen, der am besten für ihn geeignet ist.

Durch den Einsatz einer klar definierten Tarifstruktur, bestehend aus mehreren optionalen mehrteiligen Tarifen, gelingt es einem Anbieter verschiedene Marktsegmente mit unterschiedlichen Verwendungsprofilen gleichzeitig anzusprechen und die Konsumentenrente möglichst gewinnbringend abzuschöpfen [Skie99]. Jedoch ist der bloße Einsatz der Mechanismen zur mengenbezogenen Preisdifferenzierung nicht ausreichend um den wirtschaftlichen Erfolg zu garantieren. Vielmehr muss die gewählte Tarifstruktur und die Höhe der leistungsabhängigen Preise sehr sorgfältig gewählt werden; weil hierdurch sowohl die Nachfragemenge, als auch der Nutzeranteil und damit die erzielten Gewinne, signifikant beeinflusst werden, stellt die Definition der Tarifstruktur zweifelsohne eine der wichtigsten Aufgaben für den erfolgreichen Betrieb von kommerziellen internetbasierten Diensten dar [StRL01].

Entscheidend für die gewinnoptimale Ausgestaltung der Tarifstruktur ist die Kenntnis über individuelle Zahlungsbereitschaftsfunktionen der Kunden, welche die Abbildung des Nutzungsverhaltens ermöglicht. Zahlungsbereitschaftsfunktionen modellieren nicht nur Entscheidungen bzgl. "Kauf" oder "Nicht-Kauf", sondern die nachgefragte Menge in Abhängigkeit der angebotenen Tarife und die Höhe der erzielten Konsumentenrente. Es können Aussagen über Auswahlentscheidung, über die Existenz von Kannibalisierungseffekten zwischen den Tarifen untereinander sowie über Potenziale für Marktexpansionseffekte durch Einführung zusätzlicher Tarife getroffen werden.

2 Problemstellung und Zielsetzung der Arbeit

Das Dissertationsvorhaben beschäftigt sich mit der Bestimmung von Zahlungsbereitschaftsfunktionen und dem Ableiten von optimalen Tarifstrukturen im Kontext von internetbasierten Diensten. Hierzu existieren bereits verschiedene, mehrfach angewendete Verfahren (vgl. [Clem00; IyJK06]), jedoch weisen sie grundsätzlich zwei Probleme auf: Zwar stehen für die Umsetzung in der Praxis verschiedene Werkzeuge wie Befragungstools der Firma Sawtooth oder Optimierungstools von Lindo zur Verfügung; die Anwendung der Methoden und die sinnvolle Verknüpfung der Zwischenergebnisse erfordert jedoch interdisziplinäres Expertenwissen aus den Bereichen Software Engineering, Preisgestaltung, Operations Research sowie

Quantitative Methoden. Es existiert keine ganzheitliche Lösung, welche Unternehmen in sämtlichen Phasen der Tarifgestaltung – angefangen bei der Wahl der Erhebungstechnik bis zur Ermittlung optimaler Tarife (siehe Phasen in Abb. 1) – technische Unterstützung bietet. Als Folge gelingt die Tarifoptimierung nur mit hohem Aufwand bzw. es wird darauf komplett verzichtet. Mit dem Hintergrund, dass Anbieter internetbasierter Dienste in extrem dynamischen Märkten agieren und Tarifoptimierungen nicht einmalig, sondern regelmäßig durchgeführt werden sollten, besteht großes Potenzial zur Gewinnsteigerung.

Das zweite Problem betrifft die Ausgestaltung der Datenerhebung. Zugehörige Befragungen werden meist als komplex wahrgenommen – besonders, wenn ein Kunde mehr als eine Mengeneinheit zu unterschiedlichen Konditionen kaufen kann. So werden häufig Tarife vorgelegt, die sich lediglich um Nuancen in den Tarifkomponenten unterscheiden; die Wahl des optimalen Tarifs und der optimalen Nachfragemenge ist damit für den Befragten nur bedingt abschätzbar. Die Folgen sind geringe interne Konsistenz und hohe Abbruchquoten, was sich negativ auf die Aussagekraft der gewonnenen individuellen Zahlungsbereitschaftsfunktionen und folglich auf die gewinnoptimale Gestaltung der Tarifstruktur auswirkt.

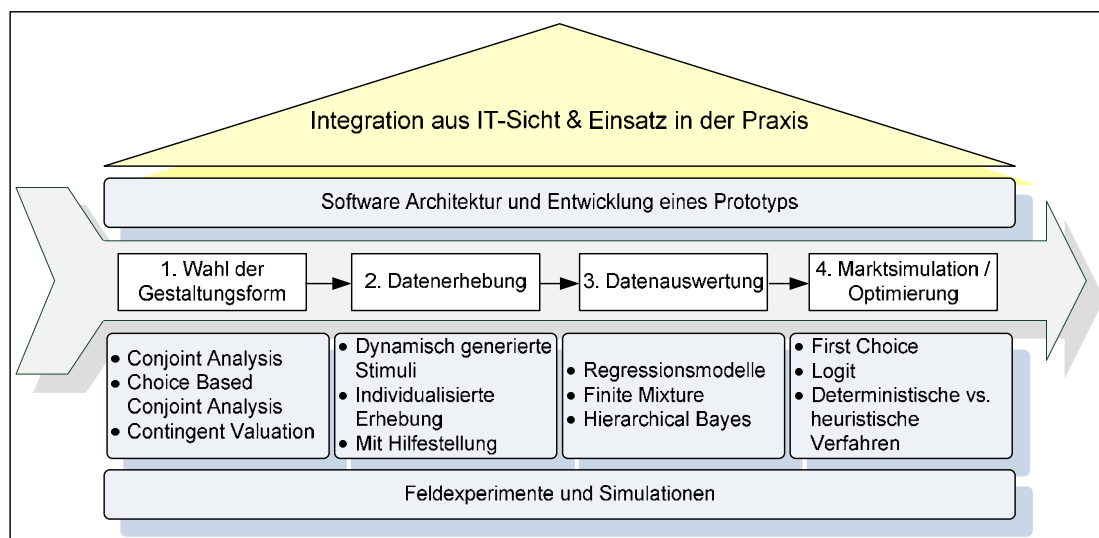


Abb. 1: Dissertationsvorhaben

Somit verfolgt die Dissertation das Ziel, Gestaltungsformen zur Erhebung individueller Zahlungsbereitschaftsfunktionen hinsichtlich der Gewinnung höherwertiger Aussagen zu untersuchen, um daraus optimale Tarifstrukturen abzuleiten. Hierbei gilt es mögliche Ausgestaltungsformen zu identifizieren und diese bezüglich der Güte der Ergebnisse zu bewerten. Verschiedene Erhebungs- und besonders Befragungsformen wurden bereits in einer Reihe von

wissenschaftlichen Arbeiten untersucht und miteinander verglichen (vgl. [ShGS05]). Jedoch wurden diese individuell auf einzelne Problemfelder ausgerichtet, so dass eine verallgemeinerte Anwendbarkeit dieser Methoden noch nicht gegeben ist.

Die Dissertation möchte nicht nur die theoretischen Grundlagen liefern, sondern darüber hinaus für Anbieter internetbasierter Dienste die Möglichkeit schaffen, die gewonnenen Erkenntnisse mit einfachen Mitteln direkt in eigene Web-Infrastrukturen zu integrieren. Daher hat die Dissertation zusätzlich die Realisierung einer flexiblen Gesamtlösung als Prototyp zum Ziel, welche sich entsprechend den unternehmensspezifischen Anforderungen konfigurieren lässt und sämtliche Phasen zur Bestimmung optimaler Tarifstrukturen unterstützt. Eine solche Anwendung kann besonders für das Management von Diensteanbietern im Internet von großer Bedeutung sein, da sie einerseits einen Überblick über die verfügbaren Instrumente und Verfahren bietet und zum anderen Möglichkeiten bereitstellt, die Effizienz der eigenen Preisstrategie gezielt analysieren und verbessern zu können.

3 Relevanz für die Wirtschaftsinformatik

Die Empfehlung, internetbasierte Dienste nicht pauschal, sondern in Abhängigkeit des tatsächlichen Nutzungsverhaltens abzurechnen, wird in der Wissenschaft seit Mitte der 90er stark diskutiert (vgl. [MaVa95; SzGM02]). Die Umsetzung dieser Empfehlungen ist aber erst seit den letzten Jahren verstärkt in der Praxis zu beobachten. Zwei bedeutende Trends in diesem Zusammenhang sind der gerade stattfindende Wertewandel auf dem Paid Content-Markt und der Technologiewandel von der klassischen Softwareanwendung zu Service-orientierten Architekturen (kurz SoA).

In der Anfangszeit des Internets hatten kommerzielle Geschäftsmodelle große Schwierigkeiten sich durchzusetzen, da bei den Nutzern im Internet eine weit verbreitete „for free“ Mentalität für digitale Güter herrschte. Erst in den letzten Jahren kam es zu einer Abkehr dieser Mentalität am Paid Content-Markt und es war eine steigende Akzeptanz bzw. Bereitschaft beobachtbar, für digitale Inhalte zu bezahlen [Ster05]. Darauf aufbauend konnten sich in der Praxis bereits verschiedene neue Geschäftsmodelle etablieren, wie beispielsweise www.xing.com (früher www.openbc.com) im Bereich Business-Netzwerke oder www.web.de im Bereich Internetportale. Diese Geschäftsmodelle bieten in der Regel kostenlose Grundleistungen, für die erweiterte

Zusatzleistungen (Premium-Dienste) gegen eine je nach Leistung gestaffelte Gebühr hinzubuchbar sind.

Der Begriff „Service-orientierte Architekturen“ repräsentiert die derzeitigen Veränderung der IT-Landschaft von Unternehmen, bei der externe Dienste als autonome Komponenten in eigene Systeme zum Einsatz kommen und mit eigenen Diensten zu neuen Geschäftsanwendungen kombiniert werden. Ein wichtiges Ziel von SoA ist die Komplexitätsreduktion [DoMZ05]. Diese wird durch die lose Koppelung der involvierten Dienste und die Abkopplung der zugrundeliegenden Technologie des Dienstenachfragers von dem Anbieter erreicht. Davon profitieren vor allem Unternehmen in Märkten, die einem ständigen Wandel unterlegen sind und den sie mit einer flexiblen Geschäftsstrategie begegnen müssen; die erforderliche Flexibilität bei der Umsetzung neuer Geschäftsanforderungen innerhalb der IT-Landschaft wird dadurch gewonnen, dass sie bestimmte Teile ihrer Geschäftsprozesse outsourcen und mit ihrer bestehenden Infrastruktur zu neuen Geschäftsanwendungen verknüpfen [StBM04]. Darüber hinaus gewinnen durch das SoA-Paradigma nun auch kleinere und mittelgroße Unternehmen Zugang zu Know How und Ressourcen, die sonst nicht für sie zur Verfügung ständen [DoMZ05]; anstatt teure Lizenzen für Softwareanwendungen zu erwerben, mieten Sie den Zugriff auf die Anwendungen. Aus Sicht der Diensteanbieter findet im Zuge dessen derzeit eine Transformation der klassischen technologieorientierten IT-Abteilung zu einem kundenorientierten IT-Dienstleister statt, der den Zugriff auf seine Dienste nicht kostenfrei zur Verfügung stellt, sondern im Rahmen seines Geschäftsmodells die Nutzung der Dienste kommerziell verwertet.

In den letzten Jahren hat sich die Wirtschaftsinformatik verstärkt damit auseinander gesetzt, wie geschäftliche Transaktionen für kostenpflichtige internetbasierte Dienste abzuwickeln sind. Hierzu gehören verschiedene Spezifikationsvorschläge zur Beschreibung kommerzieller Dienste (vgl. z.B. [OvTh05]) sowie technische Realisierungsvorschläge (vgl. z.B. [BGRH05; Fank99; Hill04]). Auch die Bestimmung der optimalen Preishöhe für internetbasierte Dienste erlangte in der Wirtschaftsinformatik eine sehr hohe Bedeutung, wobei hier unterschiedliche wissenschaftliche Strömungen existieren. Diese differenzieren sich durch die Funktion, welche die Preissetzung innerhalb der jeweiligen Geschäftsmodelle einnimmt.

In der Literatur werden folgende Funktionen des Preises unterschieden: Eine wichtige Funktion des Preises ist die Maximierung der Gewinne der Diensteanbieter [AgSe00]. Hierbei beeinflusst ein Anbieter mit dem Preis gezielt das Nutzungsverhalten, um die Differenz aus Umsatz und den dafür aufzuwendenden Kosten zu maximieren. Im Gegensatz dazu haben mikroökonomi-

sche Überlegungen eine generelle Steigerung der Wohlfahrt (sowohl für den Anbieter als für den Nachfrager) zum Ziel. So konnte beispielsweise innerhalb einer Studie eine beträchtliche Steigerung der Wohlfahrt durch Einführung von mehrteiligen Tarifen gegenüber von einfachen Pay-Per-Use Tarifen nachgewiesen werden [NgWe74]. Eine weitere Funktion des Preises ist Vermeidung von Überlastung der verfügbaren Kapazitäten (siehe „Congestion Pricing“, z.B. [ChKo03; Lin05]). Durch beispielsweise zeitliche Differenzierung wird dem Kunden ein Anreiz geboten, die Dienste hauptsächlich zu bestimmten Tageszeiten in Anspruch zu nehmen. Die unterschiedlichen Preise je Tageszeit steuern die Nachfrage und damit auch die Netzauslastung zu bestimmten Zeitpunkten, wodurch ein Anreiz zur Nutzung in weniger bevorzugten Zeiten geschaffen wird. Natürlich sind auch Kombinationen der Funktionen des Preises innerhalb eines Modells denkbar, indem mehrere Ziele in einem hierarchischen Zielsystem angeordnet werden. In der Literatur existieren bereits verschiedene Modellvorschläge, mit denen im Sinne eines Decision Support Systems die optimale Preishöhe für internetbasierte Dienste bestimmt werden kann (vgl. [AgLS02; ChKo03; Lin05]). Bei diesen Modellen werden jedoch meist mindestens eine der folgenden vereinfachenden Annahme getroffen, wodurch die Übertragung der entwickelten Modelle nur schwer in die Realität möglich ist [EsDa06]:

1. Es wird davon ausgegangen, dass einem Anbieter internetbasierter Dienste sowohl die Nachfragemenge als auch die Höhe der maximalen Zahlungsbereitschaft seiner Kunden a-priori bekannt sind (vgl. [AgLS02; Lin05]). Die Zahlungsbereitschaft ist eine Hilfsgröße aus dem Bereich Ökonomie, welche nicht direkt messbar ist. Die Erhebung der Zahlungsbereitschaft bleibt in den meisten Fällen ungeklärt und ist in der Wissenschaft immer noch eine große Herausforderung [Skie99].
2. Ferner wird häufig davon ausgegangen, dass ein Kunde genau eine Mengeneinheit konsumiert. (vgl. [AgLS02; ChKo03]) Hingegen sind internetbasierte Dienste im Gegensatz zu klassischen Softwareanwendungen keine Güter, die ein Kunde genau einmal käuflich erwirbt und anschließend beliebig häufig nutzen kann. Wie bereits am Anfang dieses Kapitels argumentiert, ist es sowohl für den Anbieter als auch für den Nachfrager attraktiv, die Dienste auf nutzungabhängiger Basis abzurechnen. Diesen Sachverhalt können diese Modelle nicht abbilden.
3. Wenn die Nutzung von mehreren Mengeneinheiten in Modellen zugelassen ist, wird unterstellt, dass der Nutzen für jede zusätzliche Mengeneinheit konstant bleibt. Alternativ wird angenommen, dass ein Kunde eine feste Nachfragemenge abnimmt, solange der vom Anbieter geforderter Preis unter der maximalen Zahlungsbereitschaft liegt (vgl. [Lin05]). Mit beiden

Vereinfachungen wird impliziert, dass die Menge der nachgefragten Einheiten unabhängig vom geforderten Preis pro Mengeneinheit ist. Jedoch hat sich gerade am Kommunikationsmarkt gezeigt, dass diese Annahme nicht zutrifft und der Nutzungspreis mit der Nachfragemenge signifikant negativ korreliert [Dola97].

4. Als eine weitere Vereinfachung werden oftmals Modelle für stark eingeschränkte Spezialfälle gelöst – beispielsweise für zwei Konsumenten und zwei Produkte (vgl. [AgLS02]). Die Verallgemeinerung in die Praxis auf eine wesentlich höhere Anzahl an Konsumenten bzw. Produkten scheitert jedoch meist aufgrund der exponentiell steigenden Komplexität.

Zusammengefasst ist festzuhalten, dass die Preisbildung bei internetbasierten Diensten – insbesondere die Bestimmung optimaler Tarife – weiterhin ein ungelöstes Problem darstellt. Die Erhebung der nicht direkt messbaren Zahlungsbereitschaftsfunktionen und der Zusammenhang zwischen Preishöhe und nachgefragten Mengeneinheiten werden in den meisten Arbeiten vernachlässigt. An dieser Stelle möchte die Dissertation einen Beitrag leisten und konzentriert sich dabei auf Szenarien, in denen die Strategie der Anbieter darin besteht, die Gewinnmaximierung durch eine zielgerichtete Preisbildung ihrer Dienste zu erreichen. Das zugrunde liegende Konzept wird im folgenden Kapitel näher erläutert.

4 Zahlungsbereitschaftsfunktionen im Kontext der Tarifoptimierung

Als zentraler Bestandteil des Dissertationsvorhabens erfolgt in diesem Kapitel eine Einführung in das Konzept von Zahlungsbereitschaftsfunktionen (im angelsächsischen Raum als "willingness-to-pay function" (kurz WTP) bezeichnet [Skie99]). Mit Zahlungsbereitschaftsfunktionen werden die Zahlungsbereitschaften für einen Dienst oder ein Gut in Abhängigkeit der nachgefragten Mengeneinheiten q beschrieben. Dass das Nutzungsverhalten stark von dem gewählten Tarif abhängt, lässt sich an einem einfachen Beispiel verdeutlichen, bei dem einem Kunde ein Pay-Per-Use-Tarif und eine Flat-Rate angeboten werden: Unter der Annahme, dass beide Tarife prinzipiell für den Kunden attraktiv sind, wird sich sein tatsächliches Nutzungsverhalten sicherlich in beiden Tarifen unterscheiden; da der Konsum einer zusätzlichen Mengeneinheit bei der Flat-Rate kostenlos ist, würde der Kunde den Dienst wesentlich häufiger in Anspruch nehmen, als innerhalb des Pay-Per-Use-Tarifs.

Über Zahlungsbereitschaftsfunktionen lässt sich das preisabhängige Nutzungsverhalten abbilden. Dabei unterstellen Zahlungsbereitschaftsfunktionen nach dem ersten Gossen'sche Gesetz,

dass der Konsum von zusätzlichen Gütereinheiten zwar einen Zusatznutzen (Grenznutzen ≥ 0) stiftet, dieser Zusatznutzen sich aber mit steigenden Mengeneinheiten q reduziert (marginaler Grenznutzen ≤ 0). Dieses Konzept wird vor allem in den Wirtschaftswissenschaften eingesetzt, um Auswirkungen von Preisänderungen auf den Markt einzuschätzen. Der Funktionsverlauf wird dabei meist in Form einer quadratischen oder einer modifiziert exponentiellen Zahlungsbereitschaftsfunktion dargestellt [Skie99]. Der Hauptunterschied beider Darstellungsarten ist, dass bei der quadratischen Zahlungsbereitschaftsfunktion im Gegensatz zur modifiziert exponentiellen eine Sättigungsmenge unterstellt wird. Dies bedeutet, dass ab einer bestimmten Mengeneinheit der zusätzliche Konsum keinen zusätzlichen Nutzen stiftet. Beispielsweise wird ein Internet Flat-Rate Nutzer nicht seine komplette Zeit im Internet verbringen, obwohl er theoretisch dazu berechtigt wäre. Zu Demonstrationszwecken sind die zugehörigen Formeln sowie der Funktionsverlauf der quadratischen Zahlungsbereitschaftsfunktion in Abb. 2 dargestellt.

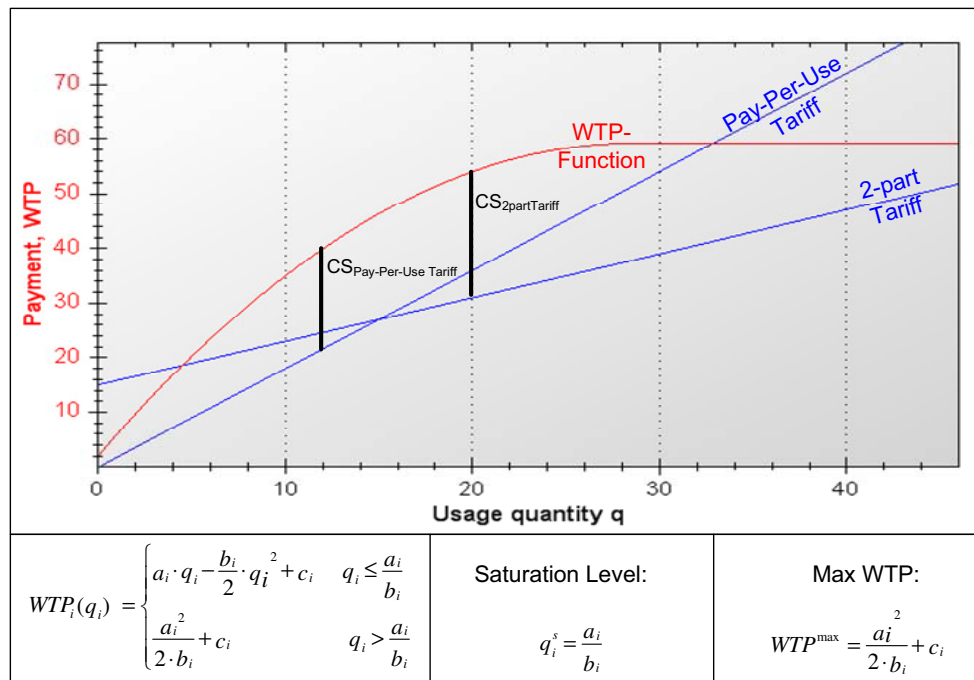


Abb. 2: Die quadratische Zahlungsbereitschaftsfunktion

Die Erhebung der Parameter a_i , b_i und c_i zur Beschreibung der individuellen Zahlungsbereitschaftsfunktion kann grundsätzlich auf drei verschiedene Arten erfolgen: Über Nutzungsdaten (bzw. Transaktionsdaten), Nutzungsangebote (z.B. Auktionen) oder Präferenzdaten. Keiner dieser drei Möglichkeiten ist eindeutig den anderen überlegen, da alle Verfahren sowohl Vor- als auch Nachteile in sich vereinen. Nutzungsdaten erfassen das tatsächliche Verhalten und können daher als valide angesehen werden. Der Nachteil besteht darin, dass die Preise in echten

Märkten nicht ausreichend variieren oder die Erhebung sich zu aufwendig bzw. kostenintensiv gestaltet [BeAk94]. Zudem ist fraglich, ob alle Konsumenten im Markt wirklich sofort auf Preisänderungen reagieren oder ob Verzögerungseffekte eintreten. Nutzungsangebote (beispielsweise erhoben durch bestimmte Auktionsformen wie Englische Auktion, BDM oder Vickrey Auktion) sind anreizkompatible Verfahren zur Offenlegung der tatsächlichen Zahlungsbereitschaft und sind deshalb in Wirtschaftswissenschaften populär [Lin05]. Der Nachteil besteht aber darin, dass diese Verfahren sehr aufwendig durchzuführen sind und es nur bedingt gelingt, individuelle Zahlungsbereitschaften in Abhängigkeit unterschiedlicher Mengeneinheiten zu ermitteln. In der Praxis fanden diese Verfahren daher nur vereinzelt Anwendung [ShGS05]. Präferenzdaten sind im Rahmen von Umfragen sehr kostengünstig zu erheben und lassen systematische Variationen zu. Die gewünschten hypothetischen Szenarien, über die man das Verhalten der Kunden analysieren möchte, lassen sich so zielgerichtet erforschen. Der Nachteil besteht darin, dass für die Befragten kein Anlass besteht, ihre tatsächlichen Zahlungsbereitschaften zu offenbaren und in manchen Situationen der Befragte einen Anreiz hat, durch strategische Antworten, die Ergebnisse der Datenerhebung zu seinen Gunsten zu beeinflussen. Das Wissen über die Zahlungsbereitschaftsfunktionen der Kunden ermöglicht einem Diensteanbieter, Aussagen über das Marktverhalten zu treffen und optimale Tarife hinsichtlich eigener Zielkriterien zu ermitteln. Das Marktverhalten kann erklärt werden, indem einzelne Tarifkomponenten geändert werden und Zu- bzw. Abwanderungsquoten, Kannibalisierungseffekte oder Preiselastizitäten bestimmt werden. Optimale Tarife lassen sich über geeignete Optimierungsmodelle berechnen, indem die Parameter der individuellen Zahlungsbereitschaftsfunktionen sowie zusätzliche Informationen als Eingabe eingesetzt werden. Beispielsweise die Gewinnoptimierung einer zweiteiligen Tarifstruktur bestehend aus einer Grundgebühr und einem Nutzungspreis könnte über folgendes Modell berechnet werden [Skie99]:

$$\pi(F_J, p_J) = \sum_{j \in J} \sum_{i \in I} \left[(p_j - k_v) \cdot q'_i(p_j) + F_j \right] \cdot z_{i,j} \rightarrow \max! \quad (1)$$

$$CS_{i,j}[q'_i(p_j)] + M \cdot (1 - z_{i,j}) > 0 \quad (i \in I, j \in J), \quad (2)$$

$$CS_{i,j}[q'_i(p_j)] + M \cdot (1 - z_{i,j}) \geq CS_{i,j'}[q'_i(p_{j'})] \quad (i \in I, j \in J, j' \in J \wedge j > j'), \quad (3)$$

$$CS_{i,j}[q'_i(p_j)] + M \cdot (1 - z_{i,j}) > CS_{i,j'}[q'_i(p_{j'})] \quad (i \in I, j \in J, j' \in J \wedge j < j'), \quad (4)$$

$$F_j \geq 0; p_j \geq k_v; z_{i,j} \in \{0,1\} \quad p_j \geq k_v \quad (i \in I, j \in J), \quad (5)$$

wobei:

- π : Gewinn,
- F_j : Grundpreis des j-ten Tarifs,
- F_J : Menge der Grundpreise aller Tarife $j \in J$,
- I : Indexmenge der Konsumenten,
- J : Indexmenge aller Tarife,
- N_j : Anzahl der Nutzer des j-ten Tarifs,
- p_j : Nutzungspreis des j-ten Tarifs,
- p_J : Menge der Nutzungspreise aller Tarife $j \in J$,
- $q'_i(\cdot)$: Nachfragefunktion des i-ten Konsumenten,
- z_{ij} : Binäre Nutzervariable zur Darstellung, ob der i-te Konsument den j-ten Tarif wählt ($z_{ij}=1$) oder nicht ($z_{ij}=0$).

Die Optimierungsverfahren basieren auf der Annahme rational handelnder Kunden, die danach streben, ihren Nettonutzen – also die Konsumentenrente – zu maximieren. Dieser Nutzen berechnet sich für eine Mengeneinheit q durch die Differenz der Zahlungsbereitschaftsfunktion und den verbundenen Ausgaben (siehe Abb. 2). Die Bestimmung gewinnoptimaler Tarife orientiert sich an ertragswirtschaftlichen Ansätzen und geschieht durch Maximierung der Produzentenrente. Hierbei können neben dem Auswahlverhalten der Kunden zusätzliche Restriktionen (z.B. durch Hardwarebeschränkungen) berücksichtigt werden.

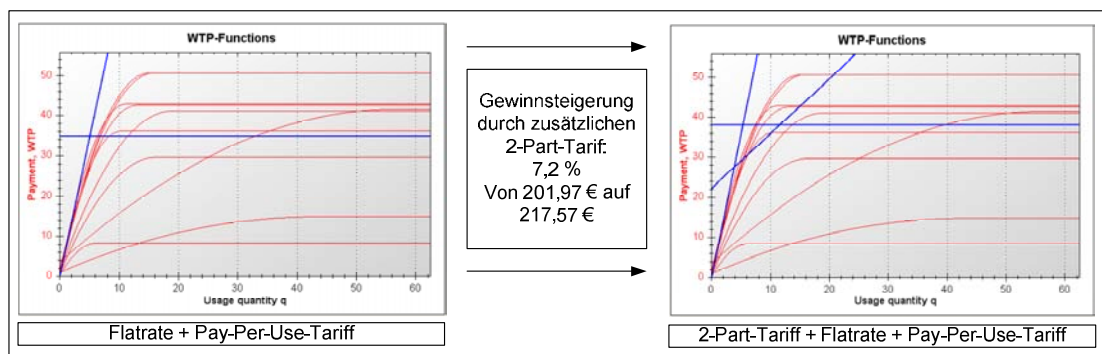


Abb. 3: Vergleich zweier Tarifstrukturen

Abb. 3 visualisiert das Ergebnis solcher Optimierungen in einem reduzierten Beispiel, in dem jeweils basierend auf 10 bekannten Zahlungsbereitschaftsfunktionen die Höhe der Flat-Rate und

des Pay-Per-Use-Tarif bestimmt wurden. Im zweiten Teil wurde ein zusätzlicher zweiteiliger Tarif in der Optimierung berücksichtigt und die Auswirkungen auf den Gewinn abgeschätzt.

5 Inhaltliche und Methodische Vorgehensweise des Dissertationsvorhabens

Die Arbeit gliedert sich in drei Teile, die interdisziplinäre Forschungsgebiete hinsichtlich des Gesamtvorhabens miteinander verbinden: Der erste Teil untersucht Strategien zur Messung von Zahlungsbereitschaftsfunktionen unter Ausnutzung dynamisch gestalteter, computergestützter Befragungstechniken. Aufbauend auf diesen Informationen beschäftigt sich der zweite Teil aus Sicht der Operations Research mit der Durchführung von Tarifoptimierung. Im letzten Teil soll aus Sicht der Wirtschaftsinformatik untersucht werden, wie die Erkenntnisse der vorherigen Teile innerhalb eines Entwurfs vereint werden können, der aufzeigt, wie im Sinne eines Decision Support Systems Managern eine stützende Hilfe bei der Auswahl optimaler Tarife geboten werden kann.

Die Ergebnisse dieser Arbeit zielen auf statische Preisgestaltungsszenarien ab, in denen die Höhe des Preises für einen Dienst vorab vom Diensteanbieter festgelegt und nicht dynamisch in Abhängigkeit der verfügbaren Ressourcen variiert wird (zum zweiten Bereich siehe die Literatur von Dynamic- bzw. Congestion Pricing, z.B. [Lin05]). Ferner konzentriert sich diese Arbeit hinsichtlich der Datenquellen auf die Erhebung von Zahlungsbereitschaftsfunktionen mittels Präferenzdaten (vgl. Kapitel 4), da sich diese besonders für neue innovative Dienste eignen, wie sie derzeit getrieben durch das SoA-Paradigma verstärkt entstehen; für diese Dienste liegen meist keine Transaktionsdaten vor, welche für die Analyse herangezogen werden könnten. Teilergebnisse dieser Arbeit, insbesondere die Auswertungen im Anschluss der Erhebung, lassen sich aber leicht innerhalb der verbleibenden Erhebungsverfahren wiederverwenden. Die offenen Forschungsfragen und die zugehörigen methodischen Vorgehensweisen werden in den folgenden Kapiteln vorgestellt.

5.1 Strategien computergestützter Erhebungstechniken

In [Iyen04; ShGS05; WoSk06] wurden bereits erfolgreich für einzelne Fallbeispiele gezeigt, wie die Erhebung und Auswertung von Zahlungsbereitschaftsfunktionen zu konzipieren ist. Die Erhebung gelingt über direkte (Contingent Valuation) oder indirekte (klassische Conjoint- und

Choice Based Conjoint-Analysen) Befragungsmethoden; für die Auswertung wurden Regressionsverfahren, Finite Mixture Modelle (zur Segmentierung) oder Hierarchische Bayes-Modelle eingesetzt. Gegenstand der Befragung waren meist Zugang zu Kommunikationsdiensten oder Zugang zu internetbasierten Diensten wie zum Beispiel kostenpflichtige Artikeln von Online-Zeitungen. Eine Empfehlung, welche der vorgeschlagenen Methoden sich im Bereich internetbasierte Dienste am ehesten eignen, existiert nicht. Daher gilt es, die Verfahren in verschiedenen Studien empirisch zu testen, um diejenigen Verfahren zu identifizieren, die sich hinsichtlich Prognosegüte, Validität und Komplexität am besten eignen.

Während der Befragung selbst besteht das Problem, dass die Qualität der Erhebungen durch die hohe Komplexität der Fragestellungen negativ beeinflusst werden kann. Eine Steigerung der Erhebungsqualität versprechen vor allem Computer Aided Survey Instruments (CAPI), die das Potenzial von computergestützten Befragungen besser nutzen können. Beispielsweise eine höhere Datenqualität kann erreicht werden, indem die Stimuli zielgerichtet auf die individuellen Interessen der Befragten ausgerichtet werden [WoSk06] oder Hilfestellungen geboten wird [WeSk02]. Zur Prüfung der Validität der gewonnenen Aussagen dienen Feldexperimente, in denen die abgewandelten Erhebungs- und Analyseformen empirisch getestet und mit den bestehenden Ergebnissen verglichen werden.

5.2 Optimierung von Tarifstrukturen

Um basierend auf den erhobenen Zahlungsbereitschaftsfunktionen optimale Tarife zu bestimmen, müssen die gewonnenen Informationen in einem normativen Modell eingesetzt und hinsichtlich einer Zielgröße optimiert werden. Jedoch wurde bereits in [AgLS02] festgestellt, dass die Komplexität der Optimierungsmodelle exponentiell mit der Anzahl an Konsumenten oder Tarifen ansteigt und deshalb eine exakte Optimierung nur schwer gelingt. Als Folge wurde in der Literatur meistens auf die Optimierung komplett verzichtet und aus dem Wissen über Zahlungsbereitschaftsfunktionen lediglich Aussagen über das Marktverhalten bei variierenden Tarifkomponenten abgeleitet. Diese Lücke soll geschlossen werden, indem ein bereits im Rahmen dieses Dissertationsvorhabens umgesetztes und leicht abgewandeltes Gradientenverfahren vorgestellt wird, dass die Bestimmung von angenäherten optimalen Ergebnissen gestattet. In diesem Verfahren kommt es zu einem Trade-Off zwischen der Genauigkeit der ermittelten Lösung und der benötigten Zeit. Um die Stärken und Schwächen des entwickelnden Ansatzes aufzudecken, soll eine Simulationsstudie das Optimierungsverfahren in unterschiedlichen, systematisch erzeugten Szenarien einsetzen.

5.3 Entwurf eines Entscheidungsstützungssystem zur Tarifoptimierung

Bisher gehörte die Ausgestaltung der Tarifstrukturen zu den Kompetenzen des Anbieters und wurde meist basierend auf implizierte Modelle (selten numerische Modelle) durchgeführt. In diesem Kapitel wird ein innovativer Ansatz vorgestellt, der computergestützte Hilfestellung bei der Bestimmung von Tarifen bietet. Hierbei sollen die Erfahrungen und Forschungsergebnissen aus dem vorherigen Teilen in den Entwurf einer Plattform einfließen, welches Empfehlungen hinsichtlich optimaler Tarifstrukturen liefert und so Marktanalysen vereinfacht. Eine solche Plattform hebt sich deutlich von den existierenden Ansätzen zu Online-Befragungen ab, da diese nicht nur die verschiedenen Formen der Datenerhebung selbst, sondern ebenfalls die Auswertung sowie die Ableitung von Handlungsempfehlungen aus den Daten in sich vereint. Neben der in Kapitel 4 eingeführte Methodik geht es in dieser Arbeit vielmehr um die zentrale Fragestellung, wie von der IT-Seite die existierenden Ansätze direkt für Anbieter internetbasierter Dienste im Sinne eines Decision Support Systems nutzbar gemacht werden können. Damit würde Anbietern eine Hilfestellung geboten und der Zugang zu den komplexen Theorien rund um die Bestimmung optimaler Tarife wesentlich erleichtert werden.

Die Realisierung soll über einen Top Down-Ansatz erfolgen, in der die Teilergebnisse der vorherigen Phasen in ein flexibles Gesamtsystem integriert werden. Architektonisch baut die zu entwickelnde Lösungen auf Überlegungen von Service-orientierten Architekturen auf [DoMZ05], in der die gesamte Vorgehensweise in einzelne Teilprozesse zerlegt und separat als lose gekoppelte Dienste verwirklicht werden. Die Zerlegung der Gesamtsoftware in mehrere Teilsysteme erlaubt es, die Komplexität der Fragestellung zu reduzieren bzw. bestimmte Teilbereiche – z.B. die Implementierung der Optimierungsalgorithmen – auszulagern [Fowl03]. Als erste grobe Einteilung für die Zerlegung bieten sich die Phasen aus Abb. 1 an, da diese jeweils grundlegend verschiedene Anforderungen an die zu entwickelnde Gesamtlösung stellen.

Abb. 4 vermittelt einen ersten Überblick über die Gesamtarchitektur. Die Kommunikation nach außen sowie innerhalb der einzelnen Komponenten erfolgt dabei über XML-basierte Nachrichten und soll über offene Standards wie SOAP, UDDI oder WSDL abgewickelt werden. Die Steuerungsebene (Process Layer) ist eine dedizierte Instanz, welche alle Aktivitäten innerhalb der Plattform steuert. Diese Ebene ist damit für die Definition sowie Konfiguration des Prozessworkflows verantwortlich. Der Execution Layer stellt die gesamte Geschäftslogik der Plattform in Form von Diensten zur Verfügung. Hierbei lassen sich die Dienste grundsätzlich in fünf Module einteilen, deren Datenaustausch durch eine Execution Engine gesteuert wird: Das erste

Modul (Configuration Services) dient zur Spezifikation der konkreten Erhebungsform. Ergebnis dieses Moduls ist ein auswertbares Umfragespezifikationsdokument in XML, welches von der Survey-Engine gelesen und ausgeführt werden kann. Die Ergebnisse der Befragung werden unter Ausnutzung des Umfragespezifikationsdokuments in eine Datenbank gespeichert, welche innerhalb Modul 3 (Data Analysis Services) ausgewertet werden kann, um die gewünschten beschreibenden Parameter über die individuellen Zahlungsbereitschaften zu ermitteln. Diese werden anschließend innerhalb von Modul 4 (Evaluation Services) und 5 (Recommendation Services) für die Optimierung und Bestimmung von Handlungsempfehlungen eingesetzt.

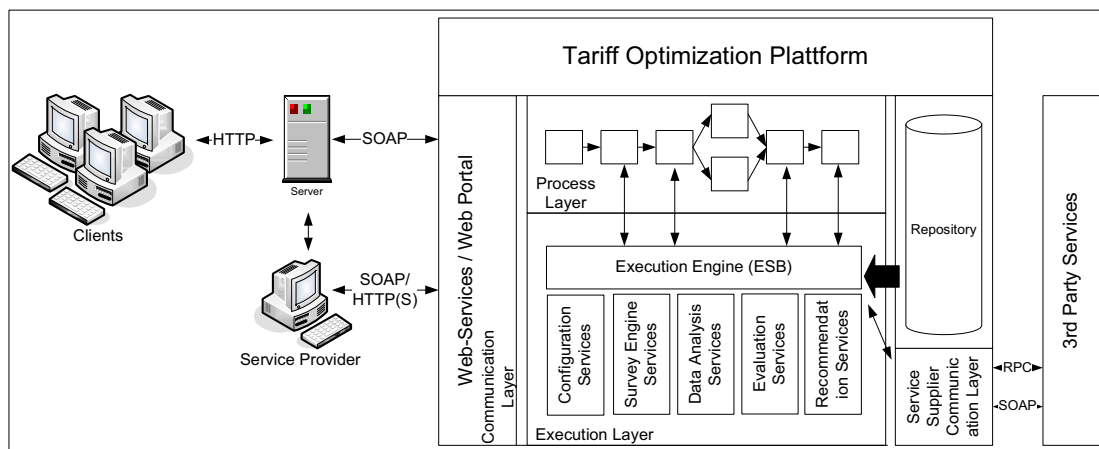


Abb. 4: Tariff Optimization Plattform

Die Plattform soll nicht komplett von Grund auf neu implementiert werden, sondern soweit wie möglich auf existierende und getestete Standardlösungen zurück greifen. Hierbei kommen 3rd Party Anwendungen externer Softwareanbieter zum Einsatz; beispielsweise wurde bereits bei der Tarifoptimierung Werkzeuge der Firma Lindo eingesetzt und ein angepasstes Gradientenverfahren als nichtlineares Optimierungsverfahren realisiert.

Die Evaluation der Plattform erfolgt durch einen Prototypen, der die Praxistauglichkeit demonstrieren soll. Dabei spielt die Identifikation und Beschreibung gradliniger Einsatzmöglichkeiten der Plattform aus Sicht der Wirtschaftsinformatik eine tragende Rolle. Besonders im Bereich SoA und/oder Grid Computing soll mit der Plattform geprüft werden, inwieweit der vorgestellte Ansatz anwendbar ist und wie die Funktionalitäten in bestehende Geschäftsanwendungen integriert werden können. In diesem Zusammenhang ist es auch wichtig die Kostenseite des Einsatzes der Plattform zu betrachten, die zur Realisierung des Potenzials zur Gewinnsteigerung notwendig ist. Die gewonnenen Erfahrungen sollen dabei künftig in Form von kontinuierlichen Weiterentwicklungen in den Prototypen einfließen.

6 Ausblick

Derzeitige Aktivitäten des Dissertationsvorhabens beschäftigen sich mit der Literaturrecherche und der Entwicklung erster Bausteine für die unterschiedlichen Phasen. Diese werden bereits im Kontext der Erhebung von Zahlungsbereitschaftsfunktionen und zur Tarifoptimierung eingesetzt. In der nicht zu fernen Zukunft (geplant August 2007) besteht das Ziel bereits einen ersten lauffähigen Prototypen mit Grundfunktionalitäten der Plattform bereitstellen zu können. Dieser soll zukünftig zuerst innerhalb von Feldstudien vorab empirisch validiert und anschließend innerhalb von Kooperation mit Unternehmen aus der Praxis eingesetzt werden. Die Abgabe der Arbeit ist für Anfang 2009 geplant.

Literaturverzeichnis

- [AgSe00] Agrawal, Vikas; Seshadri, Sridhar: Effect of Risk Aversion on Pricing and Order Quantity Decisions. *Manufacturing and Service Operations Management Journal* 4 (2000) 2, S. 410-423.
- [AgLS02] Agrawal, Vipul; Lopomo, Giuseppe; Seshadri, Sridhar: Web Based Capacity Allocation Strategies for Customers with Heterogeneous Preferences *Electronic Commerce Research* 2 (2002) 4, S. 359-384.
- [BeAk94] Ben-Akiva, Moshe; Bradley, Michael D.; Morikawa, Takayuki; Benjamin, J.; Novak, Thomas P.; Oppewal, Hamen; Rao, Vithala: Combining Revealed and Stated Preferences Data. *Marketing Letters* 5 (1994) 4, S. 335-350.
- [BGRH05] Berbner, Rainer; Grollius, Tobias; Repp, Nicolas; Heckmann, Oliver; Ortner, Erich; Steinmetz, Ralf: An Approach for the Management of Service-oriented Architecture (SoA) based Application Systems. *Enterprise Modelling and Information Systems Architectures (EMISA)*, Klagenfurt, Austria, 2005.
- [ChKo03] Cheng, Hsing Kenneth; Koehler, Gary J.: Optimal pricing policies of web-enabled application services. *Decision Support Systems* 35 (2003) 3, S. 259 - 272
- [Clem00] Clement, Michel: Interaktives Fernsehen. Analyse und Prognose seiner Nutzung. *Deutscher Universitäts-Verlag: Wiesbaden* 2000.

- [Dana02] Danaher, Peter J.: Optimal Pricing of New Subscription Services: Analysis of a Market Experiment. *Marketing Science* 21 (2002) 2, S. 119-138.
- [Dola97] Dolan, Robert J.; Simon, Hermann: *Power pricing*. Free Press: 1997.
- [DoMZ05] Dostal, W.; Jeckle, M.; Melzer, I.; Zengler, B.: *Service-orientierte Architekturen mit Web Service - Konzepte Standards, Praxis*. Spektrum Akademischer Verlag: Heidelberg et. al. 2005.
- [EsDa06] Esmaeilsabzali, Shahram; Day, Nancy A.: Online Pricing for Web Service Providers. *EDSER '06: Proceedings of the 2006 international workshop on Economics driven software engineering research*, 2006, S. 37 - 42.
- [Fank99] Fankhauser, George; Stiller, Burkhard; Plattner, Bernhard: A flexible architecture for an accounting and charging infrastructure in the Next Generation Internet. *Netnomics* 1 (1999), S. 201-223.
- [Fowl03] Fowler, Martin: *Patterns of Enterprise Application Architecture*. Addison-Wesley: Boston, MA 2003.
- [Hill04] Hill, J R: A Management Platform for Commercial Web Services *BT Technology Journal* 22 (2004) 1, S. 52-62.
- [Iyen04] Iyengar, Raghuram: *A Structural Demand Analysis for Wireless Service under Nonlinear Pricing Schemes*. Arbeitsbericht. Columbia University: New York, NY, 2004.
- [IyJK06] Iyengar, Raghuram; Jedidi, Kamel; Kohli, Rajeev: *A Conjoint Approach to Multi-Part Pricing*. Columbia University, Working Paper, 2006.
- [Lin05] Lin, Zhangxi; Ramanathan, Sathya; Zhao, Huimin: Usage-based dynamic pricing of Web services for optimizing resource allocation. *Information Systems and E-Business Management* 3 (2005) 3, S. 221 - 242.
- [MaVa95] MacKie-Mason, J.; Varian, H.: Pricing the Internet. In: Kahin, B.; Keller, J. (Hrsg): *Public Access to the Internet*. Prentice Hall: 1995, S. 269-314.
- [NgWe74] Ng, Y.; Weisser, W.: Optimal Pricing with a Budget Constraint - The Case of the Two-Part Tariff. *Review of Economic Studies* 41 (1974), S. 337-345.

- [OvTh05] Overhage, Sven; Thomas, Peter: WS-Specification: Ein Spezifikationsrahmen zur Beschreibung von Web-Services auf Basis des UDDI-Standards. Wirtschaftsinformatik 2005: eEconomy, eGovernment, eSociety. Physica (2005), S. 1539-1558.
- [ShGS05] Schulze, Timo: Optimale Nutzungspreise für Online-Zeitungen. DUV: Wiesbaden 2005.
- [Skie99] Skiera, Bernd: Mengenbezogene Preisdifferenzierung bei Dienstleistungen. Deutscher Universitäts-Verlag: Wiesbaden 1999.
- [StBM04] Steinmetz, Ralf; Berbner, R; Martinovic, I: Web Services zur Unterstützung flexibler Geschäftsprozesse in der Finanzwirtschaft. In: Sokolovsky, Z.; Löschenkohl, S. (Hrsg): Industrialisierung der Finanzwirtschaft. Gabler Wiesbaden: 2004.
- [Ster05] Stern: Akzeptanz von Paid Content im Internet steigt. 2005.
- [StRL01] Stiller, Burkhard; Reichl, Peter; Leinen, Simon: Pricing and Cost Recovery for Internet Services: Practical Review, Classification, and Application of Relevant Models Netnomics 3 (2001) 2, S. 149-171.
- [SzGM02] Szyperski, Clemens; Gruntz, Dominik; Murer, Stephan: Component Software. Addison-Wesley Professional: 2002.
- [TaCh05] Tang, Qian Candy; Cheng, Hsing Kenneth: Optimal location and pricing of web services intermediary. Decision Support Systems 40 (2005) 1, S. 129 - 141.
- [WeSk02] Wertenbroch, Klaus; Skiera, Bernd: Measuring Consumer Willingness to Pay at the Point of Purchase. Journal of Marketing Research 39 (2002) May, S. 228-241.
- [WoSk06] Wolk, Agnieszka; Skiera, Bernd: An Enhanced Method of Conjoint-Analysis to Estimate the Willingness-to-Pay for Multiple-Unit Products. Working paper, 2006.

Gestaltung kooperativer e-Financial Supply Chains

Entwicklung eines Referenzprozessmodells der e-Financial Supply Chain unter besonderer Berücksichtigung der Rolle externer Dienstleister

Thomas Krabichler

Lehrstuhl für Wirtschaftsinformatik II
Universität Regensburg
93040 Regensburg
thomas.krabichler@wiwi.uni-regensburg.de

Abstract

Die Frage nach der optimalen Gestaltung der Zahlungsabwicklungsprozesse im elektronischen Handel wurde bisher häufig auf die Auswahl geeigneter Zahlungsverfahren reduziert. Durch die Betrachtung des gesamten Prozesses von der Risikoprüfung des Kunden bis zum endgültigen Eingang der Zahlung (e-Financial Supply Chain) bieten sich jedoch neue Potenziale, Interessenskonflikte zwischen Unternehmen und Kunden zu reduzieren und die Kosten der Zahlungsabwicklung zu senken. Die Realisierung dieser Potenziale ist nicht trivial, sondern stellt für die Unternehmen eine komplexe und vielschichtige Aufgabe dar. Im Rahmen der Arbeit wird ein Referenzprozessmodell entwickelt, um vorhandenes Expertenwissen über die Gestaltung von e-Financial Supply Chains anwendungsgerecht aufzubereiten und den Unternehmen zur Verfügung zu stellen. Dabei kommt insbesondere der Identifikation und Integration von Leistungen externer Anbieter eine wichtige Rolle zu.

1 Einführung

Die jüngste Studie des Marktforschungsinstituts European Information Technology Observatory (EITO) zeigt, dass der elektronische Handel weiter an Bedeutung gewinnt. Im Jahr 2005 wurden in Deutschland Waren und Dienstleistungen im Wert von insgesamt 321 Mrd. Euro über das Internet verkauft. Bis zum Jahr 2009 sollen die Umsätze im elektronischen Handel

weiter auf 694 Mrd. Euro steigen. Der Großteil der Umsätze entfällt mit 85% auf Transaktionen zwischen Unternehmen. [Bitk06]

Es ist daher nicht verwunderlich, dass die Zahl der Unternehmen, die das Internet als Vertriebskanal nutzen, seit einigen Jahren wieder stark zunimmt [SKBW06, 117]. Als wichtigste Argumente für den Einstieg in den elektronischen Handel werden von den Unternehmen die Erschließung neuer Zielgruppen und die Bindung bestehender Kunden genannt, die zunehmend nach der Möglichkeit des elektronischen Einkaufs verlangen [SKBW06, 89]. Darüber hinaus lassen sich über das Internet auch zusätzliche Umsatzpotenziale im Ausland leichter erschließen als über herkömmliche Vertriebswege [SKBW06, 76; PeWL02, 9-17].

Um diese Potenziale des elektronischen Handels erfolgreich nutzen zu können, müssen die Abwicklungsprozesse der Unternehmen auf die besonderen Gegebenheiten im Internet-Vertrieb abgestimmt werden [ScTH05, 5; Alte02, 10]. Neben den Prozessen der Warenlieferung bzw. Dienstleistungserbringung rücken dabei zunehmend auch die Prozesse in den Fokus, die mit der Vereinnahmung der finanziellen Gegenwerte dieser Leistungen verbunden sind [Görs06, 62]. Die Aufgaben der Planung, Steuerung und Kontrolle dieser Prozesse werden im Folgenden unter dem Begriff „e-Financial Supply Chain Management“ subsumiert (vgl. Abb. 1).

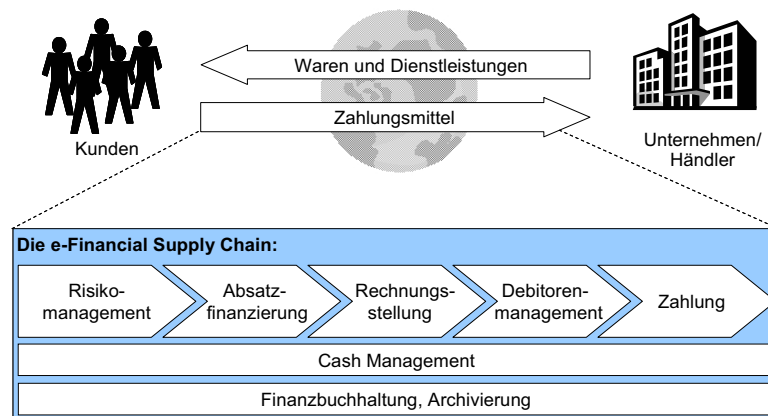


Abb. 1: Die e-Financial Supply Chain

2 Potenziale und Status des e-Financial Supply Chain Managements

Die Frage nach der optimalen Gestaltung der Zahlungsabwicklungsprozesse im elektronischen Handel wurde bisher häufig auf die Auswahl geeigneter Zahlungsverfahren reduziert [Nico04, 228-247; BKSW05; OECD 2006]. Eine Betrachtung der derzeit verfügbaren Zahlungsverfahren für den elektronischen Handel zeigt jedoch, dass keines der heute verfügbaren Zahlungsverfahren gleichzeitig die drei Anforderungen „Zahlungsgarantie“, „Verbreitung bei den Kunden“ und „durchgängige Prozessgestaltung“ der Unternehmen an die Zahlungsabwicklung ausreichend erfüllt [BKSW05, 99]. In der Folge treten bei der Zahlungsabwicklung im Internet Interessenskonflikte zwischen Unternehmen und Kunden sowie hohe Kosten für die Unternehmen auf.

Im Gegensatz zum herkömmlichen Ansatz der Betrachtung einzelner Zahlungsverfahren für den elektronischen Handel bietet die Analyse des gesamten Prozesses von der Risikoprüfung des Kunden bis zum endgültigen Eingang der Zahlung die Möglichkeit, Interessenskonflikte zwischen Unternehmen und Kunden zu reduzieren und die Kosten der Zahlungsabwicklung zu senken. Auf dieses Potenziale, den derzeitigen Status der Realisierung dieser Potenziale sowie auf die entsprechenden Hindernisse wird im Folgenden näher eingegangen.

2.1 Auflösung von Interessenskonflikten

Im Gegensatz zu stationären Handelsgeschäften ist eine Zug-um-Zug-Erfüllung bei elektronischen Handelsgeschäften nicht möglich. Von den Unternehmen werden folglich Zahlungsverfahren präferiert, die möglichst guten Schutz vor Zahlungsausfällen bieten (hierzu zählt beispielsweise die Zahlung per Vorkasse). Die Kunden nutzen stattdessen vorzugsweise Zahlungsverfahren, bei denen eine Zahlung im Falle fehlerhafter Lieferung storniert (z.B. Lastschrift oder Kreditkarte) oder zurückgehalten (Zahlung per Rechnung) werden kann. Für die Unternehmen bergen diese Zahlungsverfahren jedoch das Risiko, dass Waren von finanzschwachen Kunden oder gar in betrügerischer Absicht bestellt und anschließend nicht bezahlt werden. Auch bei der Vielzahl der speziell für den elektronischen Handel entwickelten E-Payment-Verfahren lässt sich der Zielkonflikt zwischen Unternehmen und Kunden nur unzureichend lösen. Zudem sind diese Verfahren auf Unternehmen- und Kundenseite derzeit nur wenig verbreitet [SKBW06, 136; Payp06, 90] und mit vergleichsweise hohen Kosten für den Zahlungsempfänger verbunden.

Aufgrund der gegenläufigen Interessen der Unternehmen und der Kunden besteht die Gefahr, dass vorhandene Umsatzpotenziale im elektronischen Handel nicht erschlossen werden können [OECD 2006, S. 6]. So hat eine von ibi research durchgeführte Unternehmensbefragung gezeigt, dass etwa dreißig Prozent der Unternehmen ihren Kunden ausschließlich Zahlung per Vorkasse oder Nachnahme anbieten, weitere 45% schränken die Auswahl in bestimmten Fällen, z.B. bei Neukunden oder bei Forderungen ab einer bestimmten Höhe, auf diese Verfahren ein [SKBW06, 166]. In einer Konsumentenbefragung gaben demgegenüber etwa die Hälfte der Befragten an, im Internet ungern per Vorkasse zu bezahlen; 38% der Konsumenten würde mehr online einkaufen, wenn das Bezahlen besser gelöst wäre [Payp06, 91]. Hinzu kommt, dass Unternehmenskunden, die das umsatzstärkste Kundensegment im elektronischen Handel darstellen, in den wenigsten Fällen eine Zahlung per Vorkasse akzeptieren werden.

Um den Zielkonflikt zwischen den Anforderungen der Unternehmen und denen der Kunden aufzulösen, muss die Betrachtung über die eingesetzten Zahlungsverfahren hinaus um weitere Prozessbestandteile der Zahlungsabwicklung erweitert werden, die sich auf das Zahlungsausfallrisiko auswirken. So lässt sich durch geeignete Risikoprüfungen das Risiko eines Zahlungsausfalls bei Zahlungen per Rechnung, Lastschrift oder Kreditkarte deutlich reduzieren. Beispielsweise kann unter Einbeziehung externer Quellen überprüft werden, ob die Person an der angegebenen Adresse postalisch bekannt ist, um Fehllieferungen zu vermeiden sowie die Wahrscheinlichkeit der erfolgreichen Beitreibung einer Forderung bei Nichtzahlung zu erhöhen. Zur Beurteilung der Bonität der Kunden können Risikoscorings (z.B. von Creditreform / CEG) und Verzeichnisse von Negativmerkmalen (z.B. Nichteinlösung von Lastschriften, eidesstattliche Versicherungen) herangezogen werden. Nicht zuletzt wird das Zahlungsausfallrisiko auch dadurch determiniert, wie effektiv die Prozesse im Debitorenmanagement gestaltet sind (z.B. nach welcher Frist die erste Mahnung versandt wird, ob der Kunde nur schriftlich oder auch telefonisch kontaktiert wird etc.).

2.2 Ganzheitliche Kostenbetrachtung

Die Kosten der Zahlungsabwicklung werden maßgeblich von den Transaktionsgebühren, dem manuellen Aufwand der Zahlungsabwicklung für den Unternehmen sowie die Länge der Forderungslaufzeit bestimmt. So kann der Unternehmen bei den meisten E-Payment-Verfahren relativ schnell über den finanziellen Gegenwert seiner Leistung verfügen und ist mit wenig Aufwand belastet, allerdings können die Transaktionsgebühren bei diesen Verfahren bis zu 20% der Abrechnungssumme betragen. Die Zahlung per Rechnung verursacht als Gegenbeispiel

keine direkten Transaktionsgebühren, ist jedoch mit langen Forderungslaufzeiten [SKGW04, 79] und hohen Aufwänden für die Rechnungserstellung und -versendung sowie die Zuordnung von Zahlungseingängen zu offenen Posten verbunden.

Diese manuellen Aufwände lassen sich jedoch reduzieren, indem die entsprechenden Prozesse automatisiert werden. Wie hoch die Kosten der Zahlungsabwicklung tatsächlich sind, hängt daher wesentlich auch davon ab, inwieweit Prozessschritte manuell oder automatisiert durchgeführt werden.

Wie diese Beispiele zeigen, darf sich die Entscheidung über die Form der Zahlungsabwicklung im elektronischen Handel nicht allein auf die Auswahl eines Zahlungsverfahrens beschränken. Stattdessen müssen auch die Kosten in vor- und nachgelagerten Prozessen (Risikoprüfung, Absatzfinanzierung, Rechnungsstellung, Debitorenmanagement, Cash Management, Finanzbuchhaltung und Archivierung) in die Betrachtung mit einbezogen und Verbesserungspotenziale in diesen Prozessen systematisch identifiziert und ausgeschöpft werden.

2.3 Status Quo des e-Financial Supply Chain Managements

In den e-Financial Supply Chains der deutschen Unternehmen bleiben derzeit noch erhebliche Potenziale für Prozessverbesserungen ungenutzt. Dies lässt sich anhand der Ergebnisse einer Unternehmensbefragung zu den eingesetzten Maßnahmen zur Risikoprüfung sowie des Automatisierungs- und Outsourcinggrads in der e-Financial Supply Chains feststellen.

Zur Vermeidung von Zahlungsausfällen überprüfen derzeit nur 30% der befragten Unternehmen Verzeichnisse von Negativmerkmalen, nur 14% fragen Risikoscorings ihrer Kunden ab [SKBW06, 166]. Gleichzeitig gibt der überwiegende Anteil der Unternehmen jedoch an, dass die Vermeidung von Zahlungsausfällen und das Angebot kundenfreundlicher Zahlungsverfahren für sie die größten Herausforderungen bei der Zahlungsabwicklung im elektronischen Handel darstellen [SKBW06, 160]. Insbesondere kleinere Unternehmen greifen nur selten auf die Dienstleistungen wie die Verifizierung von Lieferadressen und die Abfrage von Negativmerkmalen oder Risikoscorings zurück [SKBW06, 169]. Wie Untersuchungen in Deutschland und den USA zeigen, haben kleinere Unternehmen gleichzeitig deutlich höhere Zahlungsausfälle im elektronischen Handel hinzunehmen als größere Unternehmen [BaKH05, 88; Cybe05, 18].

Die Übertragung von Lieferdaten in das Rechnungserstellungssystem bzw. der Rechnungsdaten in das Buchhaltungssystem erfolgt bei etwa 35% der Unternehmen derzeit noch manuell

[SKBW06, 183]. Kontoauszugsdaten werden sogar von mehr als der Hälfte der befragten Unternehmen manuell in die Buchhaltungssysteme der Unternehmen übertragen [SKBW06, 183]. Dies ist vor allem deshalb problematisch, da es bei manueller Übertragung der Daten leicht zu Fehlern kommen kann, die zu Verärgerungen bei den Kunden, Verzögerungen durch Klärungs- und Abstimmungsprozesse sowie fehlender Kontrolle über die Forderungsbestände führen.

Unternehmen, für die sich eine durchgängige Automatisierung der internen Prozesse aufgrund zu geringer Transaktionszahlen nicht lohnt, bietet das Outsourcing von Teilprozessen der e-Financial Supply Chain eine überlegenswerte Alternative. So könnten beispielsweise Risikoprüfung, Rechnungsausstellung und –versand, Forderungsfinanzierung, Überwachung offener Posten oder Mahnwesen/Inkasso effizienter durch spezialisierte Dienstleister erbracht werden. Derzeit werden diese Möglichkeiten jedoch nur von wenigen der befragten Unternehmen in Anspruch genommen [SKBW06, 204]. Gleichzeitig gibt jedoch die Hälfte der Unternehmen an, die Risikoprüfung des Kunden zukünftig an externe Dienstleister vergeben zu wollen, etwa 40% wollen das Mahnwesen und Inkasso zukünftig auslagern [SKBW06, 211].

2.4 Hindernisse für die Realisierung der Potenziale

Zusammenfassend lässt sich festhalten, dass die prozessorientierte Betrachtung und Gestaltung der Zahlungsabwicklung im elektronischen Handel hohe Potenziale bietet, die Kundenzufriedenheit zu verbessern und gleichzeitig die Zahlungsausfälle und Kosten in der e-Financial Supply Chain zu senken. Dabei kommt insbesondere der Integration von Leistungen externer Anbieter eine wichtige Rolle zu.

Die Realisierung dieser Potenziale ist jedoch nicht trivial, sondern stellt für die Unternehmen eine komplexe und vielschichtige Aufgabe dar. So enthält die e-Financial Supply Chain vom Auftrag des Kunden bis zum endgültigen Zahlungseingang zum einen eine Vielzahl von möglichen Prozessschritten, zum anderen hängt der konkrete Prozessablauf z.B. von den genutzten Zahlungsverfahren, von Merkmalen der Kunden (Bestandskunde vs. Neukunde, inländischer vs. ausländischer Kunde, Verbraucher vs. Unternehmen), der Güter bzw. Dienstleistungen (physisch vs. digital, standardisiert vs. kundenindividuell) und der Transaktion ab (isolierte Transaktion vs. Transaktion innerhalb einer übergeordneten Rahmenvereinbarung). Da die e-Financial Supply Chain nicht zu den Kernprozessen eines Unternehmens zählt, kann die Verfügbarkeit der erforderlichen Kompetenzen zur Gestaltung der e-Financial Supply Chain in den Unternehmen nicht vorausgesetzt werden. Die Herausforderung besteht daher darin,

vorhandenes Expertenwissen über die Gestaltung von e-Financial Supply Chains anwendungsgerecht aufzubereiten und den Unternehmen zur Verfügung zu stellen.

3 Ziele des Dissertationsvorhabens

Zielsetzung der Arbeit ist es, die Gestaltung unternehmensspezifischer e-Financial Supply Chains durch ein Referenzprozessmodell zu unterstützen. Sie verfolgt damit eine anwendungsorientierte Zielsetzung, die nicht nur die Beschreibung von Methoden und Instrumenten des e-Financial Supply Chain Managements und die Erklärung des Beitrags dieser Methoden und Instrumente zu den Zielen des e-Financial Supply Chain Managements, sondern auch die Umsetzung der dabei gewonnenen Erkenntnisse in konkrete Gestaltungsempfehlungen umfasst.

Prozessmodelle dienen allgemein der Reduktion der Komplexität eines realen betrieblichen Ablaufs durch Modellbildung und somit der Schaffung von Transparenz [KlWi99, 192]. Prozesse werden modelliert, um sie zu verstehen, zu dokumentieren, zu analysieren und zu verbessern [OWSW03, 3; ErPe98, 6-16]. Referenzprozessmodelle sollen dabei „als Ausgangslösungen dienen, aus denen sich wirtschaftlich unternehmensindividuelle Konkretisierungen ableiten lassen“ [BHKS00, 90].

In der Literatur und der betrieblichen Praxis lässt sich eine Vielzahl von Zwecken identifizieren, die allgemein mit der Modellierung bzw. speziell mit der Referenzmodellierung verfolgt werden [Leis02, 7]. Diese Zwecke lassen sich grob in Zwecke der Organisationsgestaltung und Zwecke der Informationssystemgestaltung klassifizieren [RoSc02, 58]. Diese beiden Anwendungszwecke von Referenzprozessmodellen werden im Rahmen der Arbeit vertiefend unter dem Aspekt betrachtet, wie sich Dienstleistungen externer Anbieter in den Prozessablauf einfügen lassen. Insgesamt ergeben sich in Bezug auf Gestaltung kooperativer e-Financial Supply Chains damit vier Ziele, die mit dem zu entwickelnden Referenzprozessmodell verfolgt und die im Folgenden näher erläutert werden:

- die Unterstützung der Unternehmen bei der Gestaltung von Prozessabläufen,
- die Unterstützung der Unternehmen bei der Auswahl geeigneter Dienstleistungen,

- die Schaffung der Voraussetzungen für die effiziente Automatisierung der e-Financial Supply Chain sowie
- die Definition standardisierter elektronischer Dienstleistungen.

3.1 Unterstützung der Unternehmen bei der Gestaltung von Prozessabläufen

Die Unterstützung der Unternehmen bei der Gestaltung von Prozessabläufen entspricht dem grundsätzlichen Ziel von Referenzprozessmodellen, als Ausgangslösung bei der Gestaltung unternehmensindividueller Prozessmodelle zu dienen. Insbesondere bei kleinen und mittleren Unternehmen (KMU) können allgemein beträchtliche Informationsdefizite in Hinblick auf elektronische Handelsprozesse sowie eine mangelnde Prozesssicht festgestellt werden [ThHO00, 2]. Die Potenziale einer Umgestaltung der Prozesse und der Einbeziehung externer Dienstleister werden daher häufig gar nicht erkannt. So ergab eine Befragung von überwiegend kleinen Unternehmen, dass nur weniger als die Hälfte der Unternehmen regelmäßig prüft, ob die angebotenen Bezahlssysteme noch den Anforderungen des Unternehmens entsprechen [Payp06, 91].

Um die Unternehmen bei der Gestaltung ihrer e-Financial Supply Chain zu unterstützen, ist zu fordern, dass die Referenzprozessmodelle als Sollmodelle bei der Prozessgestaltung verwendet werden können [Broc03, 32]. Inwieweit ein bestimmtes Referenzprozessmodell tatsächlich einen „Best Practice“ repräsentiert, lässt sich jedoch nicht abschließend prüfen [Schü98, 300]. Zur Lösung dieses Dilemmas wird vorgeschlagen, Referenzprozessmodelle als Theorien zu betrachten, die nach Popper [82, 15] das wesentliche Kriterium der Falsifizierbarkeit aufweisen müssen [Schü98, 32]. Referenzmodelle sind daher um Aussagen dazu zu ergänzen, in welchem Kontext sie gültig sind und welche Ziele mit dem Referenzmodell verfolgt werden, um die Identifizierung überlegener Lösungen für den beschriebenen Kontext zu ermöglichen [Schü98, 300-304].

Neben dem Kriterium der Falsifizierbarkeit kommt dem Kriterium der Nachvollziehbarkeit eine wichtige Bedeutung zu. Popper vergleicht Theorien mit Netzen, die ausgeworfen werden, um die Realität einzufangen, sie zu rationalisieren, zu erklären und zu beherrschen [Popp82, 31]. Aufgabe des Wissenschaftlers sei es, die Maschen des Netzes immer enger zu machen [Popp82, 31]. Die im Rahmen der Arbeit zu entwickelnden Referenzprozessmodelle können daher als verfeinerte und stärker operationalisierte Formen abstrakter Theorien aus dem Finanzmanagement, Supply Chain Management und Geschäftsprozessmanagement (z.B.

Kapitalmarkttheorie, Transaktionskostentheorie, Prinzipal-Agenten-Theorie, Resource-based view) interpretiert werden. Um die Nachvollziehbarkeit der Entwicklung des Referenzprozessmodells zu gewährleisten, wird im Rahmen der Arbeit ausführlich auf die Einordnung des Financial Supply Chain Managements in das Finanzmanagement, Supply Chain Management und Geschäftsprozessmanagement sowie auf die aktuellen Herausforderungen bei der Gestaltung von Financial Supply Chains im elektronischen Handel (e-Financial Supply Chains) eingegangen.

Für eine Weiterentwicklung und Verfeinerung dieser Theorien reicht eine Beschränkung auf die bereits bestehenden Theorien jedoch nicht aus [Schü98, 12]. Vielmehr zielt die Arbeit darauf ab, in Form von Expertengesprächen, Fallstudien und quantitativen Erhebungen neue Erkenntnisse über Gestaltungsformen und -alternativen der e-Financial Supply Chain zu gewinnen. Die entwickelten theoretischen Grundlagen werden anschließend jedoch wieder dazu verwendet, die in der Praxis vorgefundenen Phänomene zu erklären und zwischen konkurrierenden Gestaltungsansätzen in der Praxis diejenigen auszuwählen, die zur Erfüllung der situationsabhängigen Ziele des Referenzprozessmodells am besten geeignet erscheinen.

3.2 Unterstützung der Auswahl geeigneter Dienstleistungen

Als weitere Zielsetzung wird mit dem Referenzprozessmodell die Unterstützung der Unternehmen bei der Auswahl geeigneter Dienstleistungen verfolgt. Im Rahmen der Arbeit werden im Wesentlichen zwei Arten von Dienstleistungen betrachtet: Dienstleistungen von Finanzintermediären und Dienstleistungen des Business Process Outsourcing.

Zu den Dienstleistungen von Finanzintermediären zählen neben den klassischen Bankdienstleistungen (Abwicklung von Überweisungen und Lastschriften, Kreditvergabe, Anlage von Liquiditätsüberschüssen) auch Dienstleistungen von Nicht-Banken, die klassische Bankdienstleistungen komplementieren bzw. substituieren. Als komplementäre Dienstleistungen können im Wesentlichen Auskunftsdienste über die Identität und Bonität des Kunden [Hamm06] und Dienstleistungen der elektronischen Rechnungsstellung und -präsentation (EBPP) durch einen intermediären Rechnungskonsolidator [SpPf06] genannt werden. Zu den substitutiven Dienstleistungen zählen unter anderem E-Payment-Verfahren für den elektronischen Handel (z.B. Firstgate, Paypal, T-Pay) und Dienstleistungen von Factoring-Instituten.

Die Auslagerung von (Teil-)Prozessen (Business Process Outsourcing) hat in den vergangenen Jahren aufgrund zunehmenden Kostendrucks und steigender Komplexität (u.a. durch die

Internationalisierung der Unternehmenstätigkeit [Heis06]) wachsende Aufmerksamkeit erfahren. Als Vorteile des Business Process Outsourcing werden unter anderem die Realisierung von Mengen- und Verbundeffekten, die Variabilisierung der Fixkosten, die Verbesserung der Liquidität, die Verbesserung der Flexibilität und Process Compliance und der Transfer operationeller Risiken genannt [DiBr04].

Gerade kleine und mittlere Unternehmen sind mit der Auswahl geeigneter Dienstleistungen jedoch häufig überfordert. So stehen in Deutschland allein über vierzig elektronische Zahlungsverfahren zur Auswahl, die sich in Bezug auf die Verbreitung bei den Kunden, das Zahlungsausfallrisiko und die Kosten stark voneinander unterscheiden [BKSW05]. Zudem ist bei der Auswahl geeigneter Dienstleistungen z.B. auch darauf zu achten, für welche Anwendungskontexte die Dienstleistungen geeignet sind, welche Mindestkontraktgrößen bei den unterschiedlichen Angeboten zu beachten sind oder wie hoch die Gefahr der Abhängigkeit von externen Dienstleistern (z.B. aufgrund mangelnder alternativer Angebote, fehlender Standards) in den jeweiligen Fällen ist.

Um die Unternehmen bei der Auswahl geeigneter Dienstleistungen zu unterstützen, wird das Referenzprozessmodell zusätzlich um Informationen zu möglichen Dienstleistungen externer Anbieter angereichert. Zum einen werden die einzelnen Teilprozessschritte mit Informationen dazu versehen, welche Arten von Dienstleistungen zur Unterstützung dieses Teilprozesses zur Verfügung stehen. Dies ermöglicht den Unternehmen die Identifizierung geeigneter Dienstleistungen und schafft eine gemeinsame Kommunikationsbasis für Unternehmen und Dienstleistungsanbieter. Zum anderen werden Checklisten dazu hinterlegt, was bei der Auswahl von Dienstleistungen zu beachten ist.

3.3 Schaffung der Voraussetzungen effizienter Automatisierung

Die Schaffung der Voraussetzungen effizienter Automatisierung betont zunächst wiederum den unternehmensinternen Anwendungsaspekt des Referenzprozessmodells. Unter dem Schlagwort „Business Process Management“ (BPM) hat die Modellierung unternehmensinterner Prozesse in jüngster Zeit wieder erheblich an Bedeutung gewonnen. Unter diesem Schlagwort werden Standards und Technologien zusammengefasst, die die effiziente Umsetzung von Prozessmodellen in lauffähige Anwendungen ermöglichen. Als Vorteile dieses Ansatzes werden die erhöhte Flexibilität und Agilität der Abläufe, die vereinfachte Prozesspflege durch Fachexperten, die verbesserte Transparenz der Prozesse und deren höhere Akzeptanz in der Organisation genannt [Stun05].

Im Gegensatz zu den klassischen Workflow-Management-Systemen basiert dieser neue Ansatz auf serviceorientierten Architekturen, also Diensten, die jeweils bestimmte Teile eines Geschäftsprozesses unterstützen (vgl. Abb. 2). Solche Dienste können zum einen Teile einer Anwendung (z.B. eines ERP-Systems) sein, zum anderen kann ein Dienst mehrere Anwendungen einschließen, die auf verteilten Servern und Datenbanken ablaufen [Stun05].

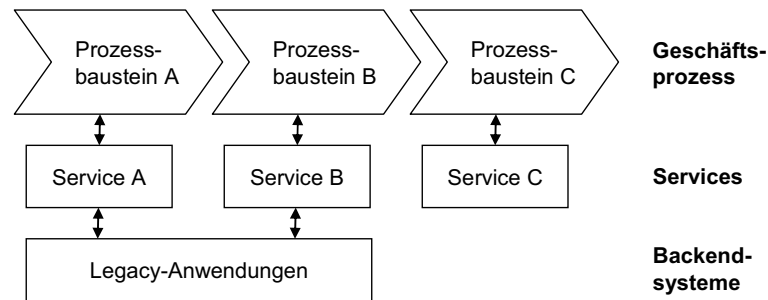


Abb. 2: Abbildung von Prozessbausteinen auf Services
Quelle: [BeMS04, 45]

Die Aufgabe der Prozessmodelle besteht darin, die Reihenfolge der Aufrufe der Dienste in formalisierter Form zu beschreiben, so dass die anschließende Ausführung des Prozesses automatisiert mithilfe von BPM-Engines erfolgen kann. Das zu entwickelnde Referenzprozessmodell soll daher auch als Vorlage für die Definition unternehmens- und BPM-Plattform-spezifischer Prozessmodelle dienen können. Im Rahmen der Arbeit wird zudem ein Vorgehensmodell dargestellt, um aus den unternehmensspezifischen Prozessmodellen lauffähige Anwendungen entwickeln zu können.

3.4 Definition standardisierter elektronischer Dienstleistungen

Eine besondere Herausforderung bei der Automatisierung von Abläufen besteht darin, auch Dienstleistungen externer Anbieter möglichst reibungslos in die Prozessabläufe zu integrieren. Sowohl auf Seite der Unternehmen als auch auf Seite der Dienstleister verursacht die Integration derzeit noch erhebliche Aufwände, die große Teile der Kooperationsvorteile wieder aufzehren [Bran03, 32]. Viele Dienstleister setzen daher bestimmte Mindestkontraktgrößen voraus, damit sich die Kosten für die Anbahnung der Geschäftsbeziehung und der individuellen Abstimmung der Dienstleistung auf die Kundensituation in absehbarer Zeit amortisieren können. Dies hat jedoch zur Folge, dass kleineren Unternehmen diese Angebote häufig verschlossen bleiben [KoPS00]. Zudem besteht die Gefahr, dass aufgrund der erheblichen

Kosten für einen Wechsel des Anbieters eine zu starke Abhängigkeit von externen Dienstleistern entsteht.

Als Lösungsansatz werden häufig die in Abschnitt 3.3 erläuterten serviceorientierten Architekturen propagiert, da es in diesem Ansatz keinen Unterschied darstellt, ob ein Dienst intern oder von einem externen Anbieter bereitgestellt wird [BeMS04]. Das Ziel besteht darin, unternehmensspezifische e-Financial Supply Chain modular aus standardisierten elektronischen Dienstleistungen zusammenzusetzen, um die Integrationskosten für Unternehmen und Dienstleistungen zu verringern und durch Verwendung der gleichen Dienste in einer Vielzahl unterschiedlicher e-Financial Supply Chains die Erzielung von Skaleneffekten zu ermöglichen. Zwar setzen sich auf technischer Ebene langsam Standards für die Realisierung serviceorientierter Architekturen durch, zur Realisierung der propagierten Potenziale reichen diese rein technischen Standards jedoch nicht aus [ReSc04, 142]. Vielmehr müssen auch auf Ebene der Prozesse Standards definiert werden, um die reibungslose Integration von modularen Prozessbausteinen zu ermöglichen. Die Entwicklung eines Referenzprozessmodells bildet hierfür den ersten Schritt, da jeder einzelne Prozessschritt des Referenzprozessmodells als modularer Prozessbaustein betrachtet werden kann. In jedem unternehmensspezifischen Prozessmodell, das auf Basis des Referenzprozessmodells erstellt wird, sollen diese Prozessbausteine möglichst ohne individuelle Anpassungen wieder verwendet werden können. Für die einzelnen Prozessbausteine werden im Rahmen der Arbeit zudem geeignete Schnittstellenbeschreibungen (Übergabe- und Rückgabeparameter; Qualitäts-, Sicherheits- und rechtliche Anforderungen [Wimm03, 122-127]) zur Verfügung gestellt.

4 Vorgehensweise zur Entwicklung des Referenzprozessmodells

Die Entwicklung eines Referenzprozessmodells der e-Financial Supply Chain stellt eine komplexe, aufwändige und durchaus auch risikoreiche Aufgabe dar. Einerseits können sich die unternehmensspezifischen Prozessmodelle, die auf Basis des Referenzprozessmodells erstellt werden sollen, in Abhängigkeit von den genutzten Zahlungsverfahren, den Merkmalen der Kunden, der Art der verkauften Güter bzw. Dienstleistungen und der Transaktion durchaus stark unterscheiden (vgl. Abschnitt 2.3). Andererseits hängt die Akzeptanz des Referenzprozessmodells durch die Unternehmen wesentlich davon ab, dass der Anpassungsbedarf des Modells möglichst gering gehalten wird. Als Ausweg aus diesem

Dilemma werden konfigurierbare Referenzprozessmodelle vorgeschlagen [BDKK02]. Konfigurierbare Referenzprozessmodelle enthalten Regeln, die besagen, wie das Referenzprozessmodell in Abhängigkeit von unternehmens- und projektspezifischen Ausprägungen von Konfigurationsparametern zu verändern ist.

Um das Risiko mangelnder Akzeptanz durch die Adressaten des Referenzprozessmodells möglichst gering zu halten, empfiehlt es sich, bei der Entwicklung des konfigurierbaren Referenzprozessmodells wiederum auf ein Referenzprozess- bzw. Referenzvorgehensmodell zurückzugreifen. Hierfür konnte ein geeigneter Ansatz identifiziert werden, der sich in fünf Phasen unterteilen lässt [BDKK02] (vgl. Abb. 3). Auf die Inhalte der einzelnen Phasen wird im Folgenden näher eingegangen (die folgenden Ausführungen basieren auf [BDKK02]).

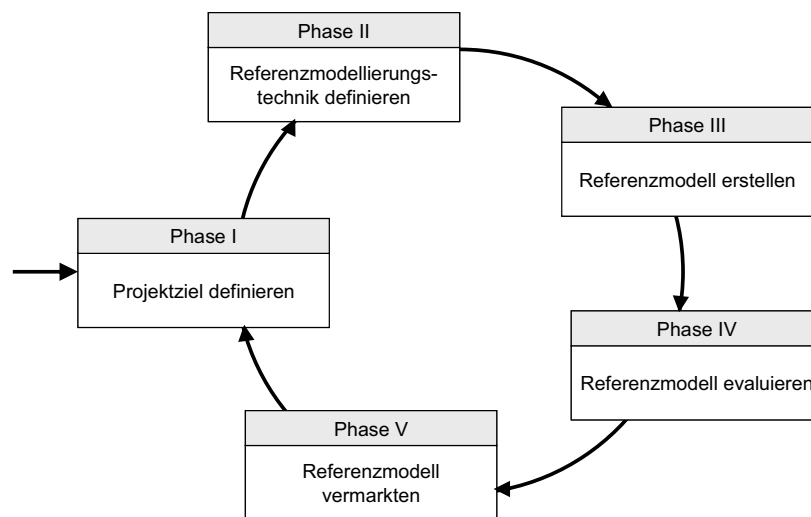


Abb. 3: Vorgehensmodell
Quelle: [BDKK02, 36]

4.1 Definition des Projektziels

Die Projektzieldefinition dient dazu, ein erstes grobes Modell des Problembereichs zu erstellen. Für die Abgrenzung des Projektziels wird vorgeschlagen, zunächst die betroffenen betrieblichen Funktionsbereiche (im Fall der e-Financial Supply Chain ist dies im Wesentlichen das betriebliche Finanzmanagement), die Unternehmensklasse sowie die Modellnutzer zu definieren, für die das Referenzprozessmodell geeignet ist. Für die Beschreibung der Unternehmensklassen und der Nutzer des Referenzprozessmodells können beispielsweise morphologische Kästen eingesetzt werden.

Aus der Projektzieldefinition lassen sich anschließend inhaltliche und methodische Anforderungen an das Referenzprozessmodell ableiten. Diese können sich beispielsweise auf die Anzahl der zu berücksichtigenden Varianten, die zu verwendende Modellierungstechnik, den Detaillierungsgrad des Modells oder die zu modellierenden Attribute (z.B. Struktur, Verhalten, Kosten, Input- und Outputdaten) beziehen. Zugleich ist die Anforderungsanalyse mit einer Analyse zu flankieren, in welchem Umfang und in welcher Form bereits Wissen über den zu modellierenden Bereich vorliegt.

4.2 Definition der Referenzmodellierungstechnik

Basierend auf der Projektzieldefinition ist zu untersuchen, welche Modellierungstechniken grundsätzlich den identifizierten methodischen Anforderungen entsprechen. Anschließend kann eine bereits vorhandene Modellierungstechnik ausgewählt, es können Modifikationen oder Neukombinationen von Modellierungstechniken vorgenommen oder sogar Modellierungstechniken vollständig neu entwickelt werden. Aus inhaltlicher Sicht wird gefordert, dass die Referenzmodellierungstechnik neben den detaillierten Prozessmodellen einen übergeordneten Ordnungsrahmen sowie geeignete Konfigurationsmechanismen enthalten muss.

4.3 Erstellung des Referenzmodells

Die Erstellung des Referenzprozessmodells sollte sowohl induktiv auf Basis bestehender unternehmensspezifischer Prozessmodelle als auch deduktiv aus allgemeinen Gestaltungsempfehlungen aus der Literatur entwickelt werden. Als Konfigurationsparameter dienen insbesondere die im Rahmen der Projektzieldefinition identifizierten Unternehmensmerkmale und Referenzmodellnutzer.

Besondere Aufmerksamkeit ist im Rahmen der Referenzmodellerstellung der Komplexität des Modells zu widmen, da sowohl die Handhabbarkeit als auch die Wirtschaftlichkeit wesentlich durch dessen Komplexität beeinflusst wird. Zur Förderung der Akzeptanz des Referenzprozessmodells ist daher darauf zu achten, dass geeignete Ansätze zur Komplexitätsreduktion bei der Modellentwicklung angemessen berücksichtigt werden.

4.4 Evaluation des Referenzmodells

Die Evaluation des Referenzmodells erfolgt zum Teil bereits während der Modellierung, soll im Folgenden jedoch als eigene Phase betrachtet werden. Die folgenden Aspekte sind in die Evaluation einzubeziehen:

- Überprüfung der Funktionsweise der Konfigurationsregeln, d.h. ob die Anwendung von Konfigurationsregeln die beabsichtigten Modellvarianten erzeugt,
- Überprüfung der Konsistenz zwischen unterschiedlichen Modellvarianten,
- Überprüfung der betriebswirtschaftlichen Gestaltungsempfehlungen und der Anforderungen an das Referenzprozessmodell.

Für die Erstellung eines Testplans wird vorgeschlagen, „einzelne Konfigurationsparameter unter ceteris-paribus-Bedingungen zu variieren oder für den Anwendungskontext repräsentative Szenarien zu bilden und diese ausführlich zu untersuchen“ [Becker/Delfmann/Knackstedt/Kuropka 2002]. Die Überprüfung der betriebswirtschaftlichen Gestaltungsempfehlungen und der Anforderungen an das Referenzprozessmodell sollte möglichst durch unabhängige Gutachter erfolgen.

4.5 Anwendung des Referenzmodells

Im Rahmen der Anwendung des Referenzmodells sollte ein möglichst enge Beziehung zu den Nutzern des Referenzmodells aufgebaut werden, um deren Erfahrungen bei der Anwendung in einen Verbesserungsprozess des Referenzmodells einfließen zu lassen. So können beispielsweise neue betriebswirtschaftliche Konzepte, neue Technologien oder Veränderungen in den Anforderungen der Nutzer dazu führen, dass das Referenzmodell angepasst werden muss. Zudem lassen sich überflüssige, fehlende oder fehlerhafte Referenzmodellteile identifizieren.

Literatur

- [Alte02] *Altenburg, Tilman: Chancen und Risiken des E-Commerce für KMU. Arbeitsbericht Nr. 9 des SEPT der Universität Leipzig. Leipzig 2002*

- [BaKH05] *Baal, Sebastian van; Krüger, Malte; Hinrichs, Jens-Werner*: Internet-Zahlungssysteme aus Sicht der Unternehmen: Ergebnisse der Umfrage IZH3. Köln 2005.
- [BDKK02] *Becker, Jörg; Delfmann, Patrick; Knackstedt, Ralf; Kuropka, Dominik*: Konfigurative Referenzmodellierung. In: Becker, Jörg; Knackstedt, Ralf (Hrsg.): Wissensmanagement mit Referenzmodellen. Heidelberg 2002, S. 25-144.
- [BHKS00] *Becker, Jörg; Holten, Roland; Knackstedt, Ralf; Schütte, Reinhard*: Referenz-Informationsmodellierung. In: Bodendorf, Grauer (Hrsg.): Verbundtagung Wirtschaftsinformatik 2000. Aachen 2000, S. 86-109.
- [BeMS04] *Berbner, Rainer; Mauthe, Andreas; Steinmetz, Ralf*: Unterstützung dynamischer eFinance-Geschäftsprozesse. In: Horster, Patrick (Hrsg.): Elektronische Geschäftsprozesse 2004. Klagenfurt 2004, S. 44-54.
- [Bran03] *Brandt, Marco*: Zahlungsabwicklung im E-Commerce. Göttingen 2003.
- [BKSW05] *Breitschaft, Markus; Krabichler, Thomas; Stahl, Ernst; Wittmann, Georg*: Sichere Zahlungsverfahren für E-Government. In: Bundesamt für Sicherheit in der Informationstechnik (Hrsg.): E-Government-Handbuch. Köln 2005.
- [Broc03] *Brocke, Jan vom*: Referenzmodellierung - Gestaltung und Verteilung von Konstruktionsprozessen. Berlin 2003.
- [Bitk06] *Bundesverband Informationswirtschaft Telekommunikation und neue Medien (Bitkom) e.V.*: Nicht mehr laufen, online kaufen! http://www.bitkom.org/de/presse/30739_39401.aspx, April 2006, Abruf am 25.07.2006.
- [Cybe05] *CyberSource*: 6th Annual Online Fraud Report. http://www.cybersource.com/resources/collateral/Resource_Center/whitepapers_and_reports/CYBS_2005_Fraud_Report.pdf, 2005, Abruf am 11.11.2005.
- [DiBr04] *Dittrich, Jörg; Braun, Marc*: Business Process Outsourcing. Stuttgart 2004.
- [ErPe00] *Eriksson, Hans-Erik; Penker, Magnus*: Business Modeling with UML - Business Patterns at Work. New York 1998.

- [Görs06] *Görs, Carsten*: (K)eine strategische Nebensache – Chancen und Risiken eines EBPP-Engagements im Banken-Umfeld. In: *Banking and Information Technology* 2/2006, S. 61-70.
- [Hamm06] *Hammel-Bonten, Christian von*: Financial Process Outsourcing. In: *Banking and Information Technology* 2/2006, S. 39-46.
- [Heis06] *Heise, Michael*: Keine Angst vor der Billigkonkurrenz. In: *Handelsblatt* Nr. 118 vom 22.06.06, S. 8.
- [KlWi99] *Klabunde, Steffen; Wittmann, Markus*: Prozeßmodellierung. In: Hofer-Alfeis, Josef (Hrsg.): *Geschäftsprozeßmanagement - innovative Ansätze für das wandlungsfähige Unternehmen*. Marburg 1999, S. 192-203.
- [KoPS00] *Kokalj, Ljuba; Paffenholz, Guido; Schröer, Evelyn*: *Zahlungsverzug und Forderungsmanagement in mittelständischen Unternehmen*. Wiesbaden 2000.
- [Leis02] *Leist, Susanne*: Bankenarchitektur des Informationszeitalters - Zielsetzung und Gestaltungsebenen. In: Leist, Susanne; Winter, Robert (Hrsg.): *Retail Banking im Informationszeitalter*. Berlin u.a. 2002, S. 3-28.
- [Nico04] *Nicolai, Sascha*: ePayments - Zahlungsverkehr entlang der eSupply Chain. In: Wannenwetsch, Helmut H.; Nicolai, Sascha (Hrsg.): *E-Supply-Chain-Management*. 2. Auflage. Wiesbaden 2004, S. 228-251.
- [OWSW03] Oesterreich, Bernd; Weiss, Christian; Schröder, Claudia; Weilkiens, Tim; Lenhard, Alexander: *Objektorientierte Geschäftsprozessmodellierung mit der UML*. Heidelberg 2003.
- [OECD06] *Organisation for Economic Co-operation and Development*: *Online Payment Systems for E-Commerce*. <http://www.oecd.org/dataoecd/37/19/36736056.pdf>, April 2006, Abruf am 14.06.2006.
- [Payp06] *PayPal/EuPD Research*: *eCommerce 2006*. <http://www.gynet-infodienst.de/img/medien/100919.pdf>, April 2006, Abruf am 29.06.2006.

- [PeWL02] *Petersen, Bent; Welch, Lawrence S.; Liesch, Peter*: The Internet And Foreign Market Expansion By Firms. Arbeitspapier Nr. 7 des Department of International Economics and Management der Copenhagen Business School. Kopenhagen 2002.
- [Popp82] *Popper, Karl R.*: Logik der Forschung. 7. Auflage. Tübingen 1982.
- [ReSc04] *Read, Cedric; Scheuermann, Hans-Dieter*: CFO - Die integrative Kraft im Unternehmen. Bonn 2004.
- [RoSc97] *Rosemann, Michael; Schütte, Reinhard*: Grundsätze ordnungsmäßiger Referenzmodellierung. In: Becker, Jörg; Rosemann, Michael; Schütte, Reinhard (Hrsg.): Entwicklungsstand und Entwicklungsperspektiven der Referenzmodellierung. Münster 1997, S. 16-33.
- [RoSc02] *Rosemann, Michael; Schwegmann, Ansgar*: Vorbereitung der Prozessmodellierung. In: Becker, Jörg; Kugeler, Martin; Rosemann, Michael (Hrsg.): Prozessmanagement. Ein Leitfaden zur prozessorientierten Organisationsgestaltung. 3. Aufl., Berlin et al. 2002.
- [ScTH05] *Schinzer, Heiko; Thome, Rainer; Hepp, Martin*: Electronic Commerce: Ertragsorientierte Integration und Automatisierung. In: Thome, Rainer; Schinzer, Heiko; Hepp, Martin (Hrsg.): Electronic Commerce und Electronic Business – Mehrwert durch Integration und Automation. 3. Auflage. München 2005, S. 1-28.
- [Schü98] *Schütte, Reinhard*: Grundsätze ordnungsmäßiger Referenzmodellierung - Konstruktion konfigurations- und anpassungsorientierter Modelle. Wiesbaden 1998.
- [SKGW04] *Skiera, Bernd; König, Wolfgang; Gensler, Sonja; Weitzel, Tim; Beimborn, Daniel; Blumenberg, Stefan; Franke, Jochen; Pfaff, Donovan*: Financial-Chain-Management: Prozessanalyse, Effizienzpotenziale und Outsourcing. Norderstedt 2004.

- [SKBW06] *Stahl, Ernst; Krabichler, Thomas; Breitschaft, Markus; Wittmann, Georg:* Zahlungsabwicklung im Internet – Bedeutung, Status-quo und zukünftige Herausforderungen. Regensburg 2006.
- [SpPf01] *Spann, Martin; Pfaff, Donovan:* Electronic Bill Presentment and Payment (EBPP). In: Die Betriebswirtschaft 61 (2001) 4, S. 509-512.
- [Stun05] *Stundzig, Steffen:* Konvergenz zwischen Model Driven Architecture und Business Process Management - Ein Ausblick. In: Fähnrich, Klaus-Peter; Thränert, Maik; Wetzel, Peter (Hrsg.): Umsetzung von kooperativen Geschäftsprozessen auf eine internetbasierte IT-Struktur - Arbeiten aus dem Forschungsvorhaben Integration Engineering. Leipzig 2005, S. 155-164.
- [ThHO00] *Thome, Rainer; Hennig, Andreas; Ollmert, Clemens:* Kategorisierung von eC-Geschäftsprozessen zur Identifikation geeigneter eC-Komponenten für die organisierte Integration von Standardanwendungssoftware. FORWIN-Bericht Nr. FWN-2000-011. Nürnberg 2000.
- [Wimm03] *Wimmer, Andreas:* Wertschöpfungsnetzwerke und deren Umsetzung in der Finanzwirtschaft - Konzepte zur unternehmensübergreifenden Integration von Produkten, Prozessen und Anwendungssystemen. Regensburg 2003.

A Field Experiment on Monetary Incentives in Prediction Markets

Stefan Luckner

Institute of Information Systems and Management (IISM)
Universität Karlsruhe (TH)
76131 Karlsruhe
Stefan.Luckner@iism.uni-karlsruhe.de

Abstract

The results of recent studies on prediction markets are encouraging. Prior experience demonstrates that markets with different incentive schemes predicted uncertain future events at a remarkable accuracy. In this paper, we study the impact of different monetary incentives on the prediction accuracy in a field experiment. In order to do so, we compare three groups of users, corresponding to three treatments with different incentive schemes, in a prediction market for the FIFA World Cup 2006. Somewhat surprisingly, our results show that performance-compatible payment does not necessarily increase the prediction accuracy.

1 Introduction

Prediction markets are a promising approach for forecasting uncertain future events. The basic idea of a prediction market is to trade virtual stocks promising certain payoffs that depend on uncertain future events. Examples comprise the outcome of an election or the results of a sports event. The Iowa Electronic Market (IEM) for predicting the outcome of the presidential elections in 1988 was the first political stock market [FNNW92]. Since then, political stock markets have been widely used as an alternative to polls and initially seemed to be the miracle cure in psephology. Apart from political stock markets, the idea behind prediction markets has also been used in various settings like in market research or business forecasting in general [SpSk03; SpSk04]. Lately, forecasting markets are also used in order to predict the outcome of sports events [LuWS06].

The principle idea is that according to the efficient market hypothesis [Fama70], prices of traded assets reflect all available information and, thus, asset prices can be used to predict the

likelihood of uncertain events. Consider a share that promises a payment of one currency unit for every percentage point a party obtains at an election. If, for example, a party wins 40 percent at the election, the participants receive 40 currency units for each share of that party they have in their portfolio. An investor who believes that the party will obtain 40 percentage points might sell his shares of this party for prices above and buy additional shares for prices below 40 currency units. Thus, the market prices reflect the expectations of the traders regarding the outcome of the election [Mans06]. Several studies have shown that the market prices of the different shares prior to the election are very close to the percentage points the respective parties win at the actual election.

The focal point of this work is to study the impact of different incentive schemes on the prediction accuracy in a field experiment. We want to elaborate on the question whether prediction markets with performance-related incentives perform better than markets with fixed payments. Somewhat surprisingly, our results show that performance-compatible incentives do not necessarily increase the prediction accuracy. Based on our results we will give advice on engineering incentive schemes for future prediction markets.

The remainder of the paper is structured as follows: The next section describes some related work on incentives schemes in the area of experimental economics and two field experiments on real-money vs. play-money prediction markets. In section 3, we then describe the setup of our field experiment we conducted during the FIFA World Cup 2006 in Germany. Furthermore, we discuss our results concerning the impact of different incentive schemes on the prediction accuracy in section 4. Thereby, we also speculate why using performance-related incentives could possibly lead to a decrease in prediction accuracy. In section 5, we finally summarize our findings and give an outlook on possible implications these results might have on designing incentive schemes for prediction markets.

2 Related Work

Previous research in the field of prediction markets has shown that play-money as well as real-money markets can predict future events at a remarkable accuracy [FNNW92; SpSk03]. So far, market operators have employed various kinds of incentive schemes in order to motivate people to take part in such markets and to reveal their expectations. Typical examples are prizes for the top performers of a market, lotteries among all traders, rankings published on the Internet or

even real-money exchanges. We suspect that the embodiment of the incentive mechanisms has a huge impact on the market quality and the prediction accuracy. Despite this, we are aware of merely two papers studying incentives for prediction markets by comparing real-money and play-money markets.

In one of these two earlier studies, Servan-Schreiber et al. found that there was no statistically significant difference between the real-money market TradeSports and the play-money market NewsFutures [SWPG04]. Rosenbloom et al., however, found TradeSports to be significantly more accurate than NewsFutures for non-sports events [RoNo06]. In case of NFL games, they produced conclusions consistent with those from Servan-Schreiber et al. Considering both studies, we believe that the impact of real-money vs. play-money at least remains an open question in the field of prediction markets. Moreover, there exists far more than one design option only for play-money markets – and most probably also for real-money markets. The strength of both studies is the large data set from real-world online experiments that both papers rely on. However, both studies do not consider any other differences apart from the use of real-money or play-money in their comparison of the two markets. Although the markets they compare are quite similar, they are by far not identical. We agree that a key difference between the two markets is that one uses real-money while the other does not. But how did some other aspects influence the prediction accuracy? It remains an open question how e.g. the number of traders and their trading activity influences the market. This seems to be an interesting question, since the number of traders per contract was not available for TradeSports. What is more, TradeSports does also levy a small fee on each transaction. How does this impact the trading behavior and the resulting share prices? The two markets – TradeSports and NewsFutures – were not identical and we thus claim that other influencing factors might have caused the results described by Servan-Schreiber et al. and also by Rosenbloom et al.

As already mentioned before, these two are the only papers dealing with incentive schemes that we are aware of in the field of prediction markets. In experimental economics however, there is quite a lot of research concerning payment schemes for participants in lab experiments. Many experimental economists most probably would insist that monetary risk is required in order to obtain valid conclusions about economic behavior. Payments based on the participants' performance are usually intended to provide incentives for rational – or at least well considered – decision making. On the other hand, there is evidence that monetary incentives do not necessarily increase performance [GnRu00]. All in all, we consider studying the impact of

different incentive schemes on the prediction accuracy of markets an open and interesting question.

3 Experimental Setup

In this section we describe the setup of the field experiment we conducted during the FIFA World Cup 2006 in Germany. Firstly, we present the basic setup. Secondly, we elaborate on the three payment schemes we studied in our field experiment and explain why we chose these three incentive schemes. Thirdly, we discuss our expected results for this experiment.

3.1 Basic Setup

In our field experiment we were operating 20 prediction markets for the last 20 matches of the FIFA World Cup 2006. As assets we traded the possible outcomes of all the matches. There were three uncertain events for every match – either team A won or team B won or there was a draw after the second half. We introduced the third asset (“draw”) although there were no draws possible in the tournament. The reason was that we did not want to consider penalty shootouts because we considered their outcome more or less unpredictable. The asset corresponding to the events that actually occurred during the World Cup was valued at 100 currency unit after the match; the other two assets were worthless.

All the markets opened about two days before the corresponding match and closed at the end of the match. As a trading platform we used the system that is currently available at www.stoccer.com. A screenshot of the web interface is depicted in Figure 1. For more information on the system itself please refer to [LuKW05].

3.2 Incentive Schemes

In total, 60 undergraduate students from the University of Karlsruhe, Germany, were taking part in our field experiment in June and July 2006. We split them into three groups of 20 students each. At the end of the FIFA World Cup the users were paid according to their group’s incentive scheme. We can thus study the impact of three different monetary incentives by comparing the prediction accuracy of the three groups of users, corresponding to three treatments with different incentive schemes. The subjects of the first group were paid a fixed amount of 50 Euro (incentive scheme 1, from now on referred to as IS₁). To subjects in the

second group we promised what we called a performance-compatible payment, also with an average amount of 50 Euro (IS_2). Performance-compatible means that the payment linearly depended on the users' deposit value in the prediction market (deposit value divided by 10.000 currency units). In the third group, individuals were paid according to their ordinal rank (rank-order tournament, IS_3). The user ranked first was paid 500 Euro, the second 300 Euro and the third 200 Euro. All the other users in this group did not receive any payment at all. This also results in an average payment of 50 Euro per person.

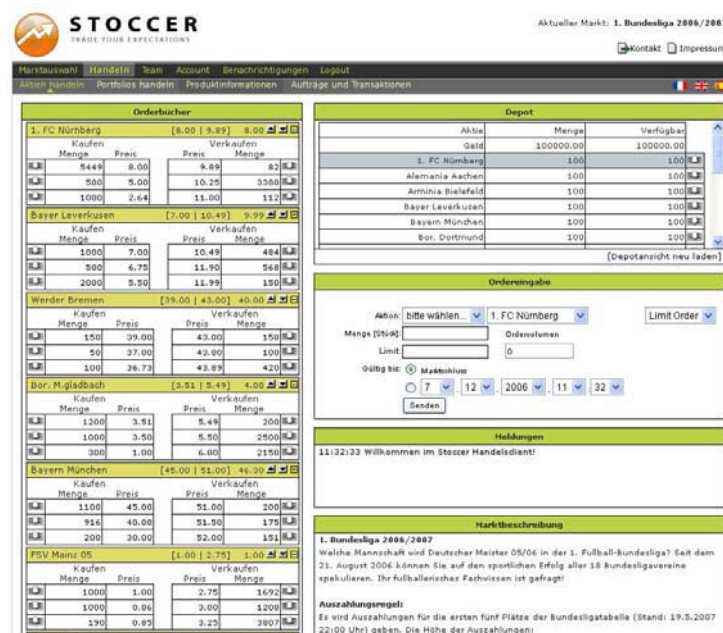


Figure 1: Web interface of the STOCER trading platform

We chose these three incentive schemes because we think they are somewhat related – although they are not the same – to incentives that we can nowadays typically observe in prediction markets, namely markets without any payment, real-money markets, and markets with rank-order tournaments.

For every group we ran the 20 separate markets on 20 soccer matches that were described in Section 3.1. Since we did not want to pay students that were not trading at all we imposed a relatively small minimum trading volume per week on all of the users. Especially in case of the first group with the fixed payment we were worried that the students might otherwise consider not to trade at all.

3.3 Expected Results

Before conducting our field experiment we expected the third group with the performance-compatible payment to be the best and the first group with a fixed payment to be the worst in terms of prediction accuracy. Let us explain the intuition behind these expectations. For members of the first group, there exists no extrinsic motivation to reveal their expectations or to be among top performers of the group. In addition, there is no incentive for them to trade more than the minimum required trading volume per week.

Members of the third group, on the other hand, receive a performance-compatible payment, meaning that every transaction directly influences their payment. Traders should consequently be motivated and try their best. Besides, traders don't want to lose money and will therefore consider very carefully what and how to trade. In short, traders with IS_3 have to "put their money where their mouth is" [Hans99].

For the second group we expected a result somewhere in between the other two groups. On the one hand, traders have a strong incentive to be among top 3 traders of the group because they will not receive any payments otherwise. This should lead to a rather high trading activity. On the other hand, the rank-order tournament provides an incentive to take higher risk compared to traders e.g. in IS_3 . Also, traders might start betting on unlikely events because they consider this the best or maybe even only way to outperform their competitors from the same group.

4 Results

In this section we will now discuss the – at first sight – probably somewhat surprising results from our field experiment. We will first compare the distribution of asset prices in the three treatments before discussing the impact of the three incentive schemes on the prediction accuracy.

4.1 Market Prices

In total, every group traded 60 assets in 20 different markets (three assets per market). In Figure 2 we can see how many assets were traded within a certain price range in each of the three treatments. The very first column for example means that 32% of the assets were traded at prices between 0 and 20 currency units in the first treatment with a fixed payment.

When comparing the three treatments we can observe that a relatively high number of assets are traded at prices between 60 and 100 currency units in the second treatment. This is exactly what

we expected because people are obviously willing to take the risk to buy assets even at rather high prices. Students in the third group with the performance-compatible payment, in contrast, do not trade any asset at a price between 80 and 100 currency units and almost not asset in the range from 60 to 80. Obviously, traders with IS₃ are not willing to take the risk of buying assets at such high prices although there is no reason why their expectations should differ that much from the traders' expectations in the other two treatments.

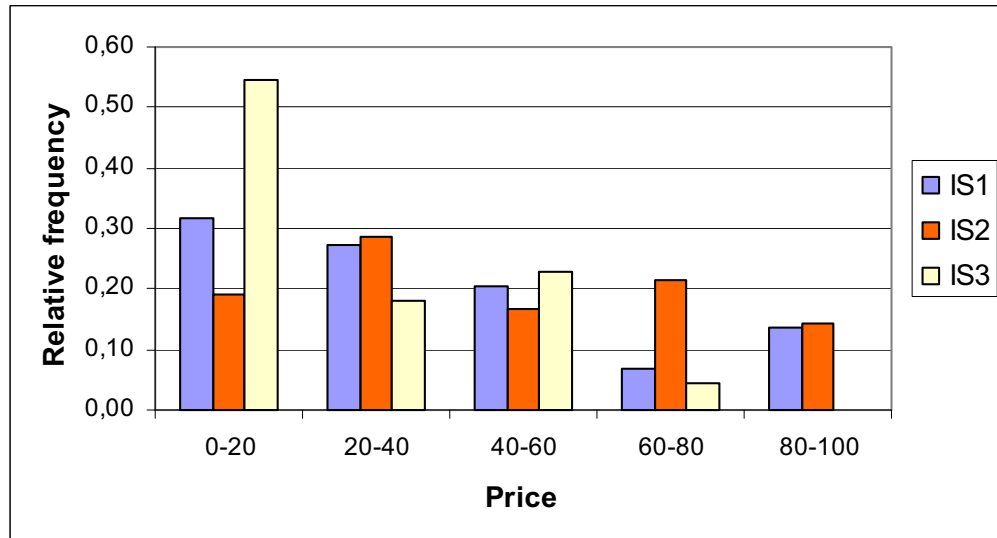


Figure 2: Distribution of asset prices in the three treatments

One again, “people are typically willing to pay less for almost anything if the money is real than if it is hypothetical” [Read05]. One explanation for this behavior of traders in the third treatment could be their risk aversion.

4.2 Prediction Accuracy

Overall, 35% of the assets with the highest share price out of the three assets per match actually corresponded to the observed outcome in case of the fixed payment and the average pre-game trading price of the asset corresponding to the outcome was 40.83 currency units. In the rank-order tournament, the favorite outcome according to the asset prices actually occurred in 45% of the cases and the average pre-game trading price of the asset corresponding to the outcome was 51.65 currency units. Finally, in case of the performance-compatible payment, the favorite outcome according to the asset prices actually occurred in merely 20% of the cases and the average pre-game trading price of the asset corresponding to the outcome was 26.64 currency units. This means, when interpreting the asset prices as probabilities the third treatment

predicted the outcome of a match worse than randomly drawing one of the three possible events. This was indeed rather surprising to us, especially since especially the rank-order tournament seems to work quite well.

However, in Section 4.1 we have already learned that asset prices seemed to be rather small in case of the performance-compatible payment. This can also be seen when calculating the sum of the three asset prices corresponding to the three possible outcomes of a match. These prices should sum up to about 100 currency units since the probability that one of the three events occurs is 100%. In case of the performance-related incentive scheme the average price of such a so called portfolio is only 53.30 currency units while it is indeed very close to 100 in the other two treatments.

To analyze the correlation between asset prices and outcome frequency in more detail, we sorted the data into buckets by assigning all of the assets to one of three price ranges according to their pre-game trading price. The size of the circles and triangles indicates how many assets prices fell into the price range. The larger the circle or triangle is, the more assets were assigned to this bucket. Figure 3 plots the relative frequency of outcome against the prices observed before the match started.

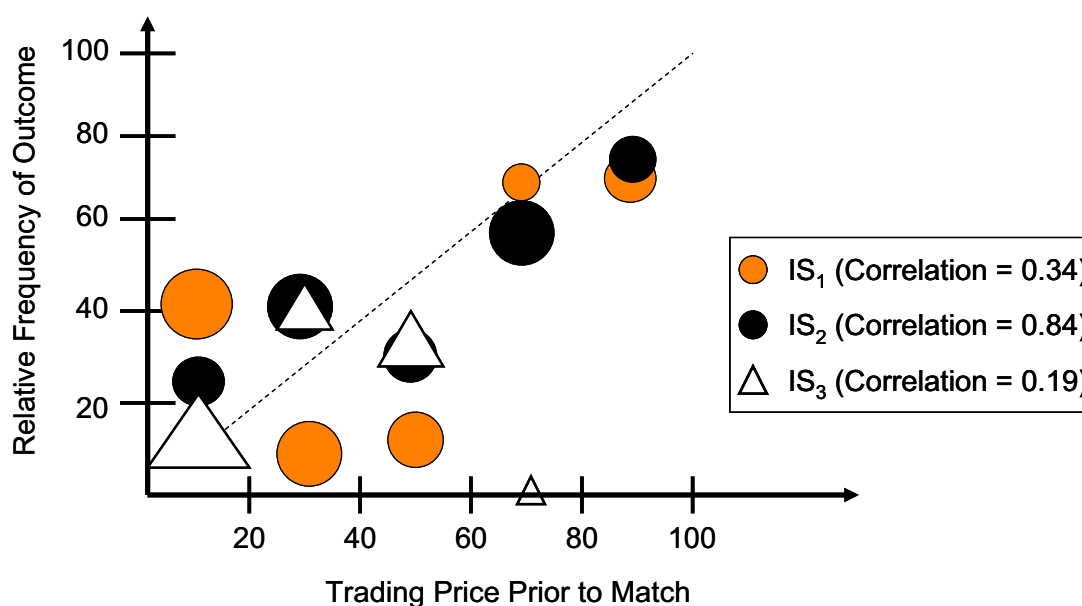


Figure 3: Market forecast probability and actual probability

For the rank-order tournament (black circles) the correlation coefficient is 0.84, while it is only 0.34 for the fixed payment and with 0.19 even worse for the performance-compatible incentive

scheme. Thus, the prediction accuracy is – in contrast to our expected results – quite poor in the third treatment IS₃.

We can now only speculate about possible reasons for this result. Besides extrinsic motivation traders might also be intrinsically motivated. This could also help to explain why even the fixed payment scheme seems to work to some extent. However, we think that the risk aversion of the traders is most likely the main reason for our results. In case of the fixed payment, traders can neither win nor lose money and risk aversion does as a consequence not matter. Moreover, traders will take quite a lot of risk in the rank-order tournament because they have to be among the top performers within their group to receive the relatively high payment. Only in the third treatment, the performance-compatible incentive scheme, traders receive an endowment of 50 Euro and could potentially lose money with every transaction they make. As a result, buyers are obviously very careful and not willing to spend too much money on any asset. Sellers on the other hand are probably willing to sell at rather low prices to avoid the risk of holding shares of an event that does in the end not occur. Maybe there would be almost no transactions of traders would not have to achieve the minimum transaction volume.

5 Summary

In this paper we have analyzed the impact of various incentive schemes on the accuracy of prediction markets. The results from our field experiment show that despite our first intuition incentive-compatible payment schemes seem to perform worse than fixed payments and rank-order tournament. Due to the risk aversion of traders, the competitive environment in case of the rank-order tournament seems to lead to the best results.

But what are the implications for designing future prediction markets? We argued in this paper that incentive-compatible payment schemes are somewhat similar to real-money markets. But can we now draw the conclusion that play-money markets will outperform real-money markets although the latter raise numerous legal and technical difficulties? We would rather be careful when answering this question based on our results because the situation might be somewhat different in prediction markets that are open to the public. In this case, there is a self-selection of traders and we would thus expect many risk-seeking traders in such a real-money market. In such a situation a performance-related payment scheme might produce much better predictions than in our field experiment.

References

- [Fama70] Fama, E. F.: Efficient capital markets: A review of theory and empirical work. In: *Journal of Finance* 25. 1970. S. 383-417.
- [FNNW92] Forsythe, R.; Nelson, F.; Neumann, G.; Wright, J.: Anatomy of an Experimental Political Stock Market. In: *American Economic Review* 82. 1992. S. 1142-1161.
- [GnRu00] Gneezy, U.; Rustichini, A.: Pay Enough Or Don'T Pay At All. In: *The Quarterly Journal of Economics* 115(3). 2000. S. 791-810.
- [Hans99] Hanson, R.: Decision Markets. In: *IEEE Intelligent Systems* 14(3). 1999. S. 16-19.
- [LuKW05] Luckner, S.; Kratzer, F.; Weinhardt, C.: STOCER - A Forecasting Market for the FIFA World Cup 2006. 4th Workshop on e-Business (WeB 2005), Las Vegas, USA.
- [LuWS06] Luckner, S.; Weinhardt, C.; Studer, R.: Predictive Power of Markets: A Comparison of Two Sports Forecasting Exchanges. In: *Information Management and Market Engineering*. T. Dreier; R. Studer; C. Weinhardt. Universitätsverlag Karlsruhe: Karlsruhe, 2006,
- [Mans06] Manski, C. F.: Interpreting the Predictions of Prediction Markets. In: *Economics Letters* 91(3). 2006. S. 425-429.
- [Read05] Read, D.: Monetary incentives, what are they good for? In: *Journal of Economic Methodology* 12(2). 2005. S. 265-276.
- [RoNo06] Rosenbloom, E. S.; Notz, W. W.: Statistical Tests of Real-Money versus Play-Money Prediction Markets. In: *Electronic Markets - The International Journal* 16(1). 2006. S.
- [SWPG04] Servan-Schreiber, E.; Wolfers, J.; Pennock, D.; Galebach, B.: Prediction Markets: Does Money Matter? In: *Electronic Markets - The International Journal* 14(13). 2004. S.

- [SpSk03] Spann, M.; Skiera, B.: Internet-Based Virtual Stock Markets for Business Forecasting. In: Management Science 49. 2003. S. 1310-1326.
- [SpSk04] Spann, M.; Skiera, B.: Einsatzmöglichkeiten virtueller Börsen in der Marktforschung. In: Zeitschrift für Betriebswirtschaft (ZfB) 74 (EH2). 2004. S. 25-48.

Zum Einsatz selbstorganisierender Koordinationsmechanismen unter Nutzung von UC-Technologien

Titus Faupel

Institut für Informatik und Gesellschaft; Abtl.: Telematik

Albert-Ludwigs Universität Freiburg
79104 Freiburg

titus.faupel@iig.uni-freiburg.de

1 Einleitung

Aufgrund der ständig steigenden Komplexität der IT in Unternehmen wird in den letzten Jahren der Selbstorganisation von betrieblichen Prozessen eine immer wichtigere Rolle eingeräumt. Die gegenwärtig entstehenden Ubiquitous Computing (UC) – Technologien werden dabei als Mittel empfunden, diese Selbstorganisation zu vertretbaren Kosten zu realisieren. Mit ihrem Einsatz und der Nutzung des theoretischen Potenzials sind allerdings noch erhebliche Probleme verbunden. So ist es nicht möglich das Verhalten und das Koordinationsergebnis eines sich selbst organisierenden Prozess zu beschreiben bzw. vorhersagen. Somit wird eine Abstimmung mit weiteren innerbetrieblichen Prozessen oder eine Anbindung an die Gesamtstrategie unmöglich. Folglich kann nicht gewährleistet werden, dass das Koordinationsergebnis zu jeder Zeit als wünschenswerte Ausprägung im Hinblick auf die Unternehmensziele angesehen werden kann. Aus diesem Grund, soll in dem Dissertationsvorhaben eine Mechanismus zur Steuerung sich selbst organisierende Prozesse entwickelt werden. Der Mechanismus soll zum einen die Vorteile einer Selbstorganisation sicherstellen und zum anderen eine gesamtzielorientierte Ausrichtung ermöglichen.

2 Ausgangslage: *Ubiquitous Computing* zur Selbstorganisation von Prozessen

2.1 Möglichkeit der Selbstorganisation durch Ubiquitous Computing

Gegenwärtig entstehen Rechner, die Weisers Vision von der unsichtbaren Technik [Weis91] Wirklichkeit werden lassen. Obwohl nach dem Gesetz von Moore und Bell damit noch keine neue Rechnerklasse entstanden ist, zeigt doch die geschichtliche Entwicklung, dass die Hardware den Anwendungen und den organisatorischen Konzepten voraus geht [Muel06]. Nicht mehr die Zentralrechner der sechziger Jahre oder die PCs mit dem Bild eines Dienstleistungen nachfragenden Anwenders (client-server), sondern die Informatisierung der Umwelt und aller ihrer Gegenstände durch eingebettete Dienste wird Realität (Ubiquitous Computing) [Muel03].

Ubiquitous Computing umfasst drei Eigenschaften, die in der Kombination die Ausgangslage für ein sich selbst organisierenden Prozess geben [LyYo02]. Erstens wird durch die Möglichkeit der Funkverbindungen und Miniaturisierung eine Mobilität erreicht, die eine freie Bewegung von Endgeräten bei Beibehaltung der gewohnten Dienste ermöglicht (Mobile Computing). Zweitens ermöglichen Sensoren und Aktoren eine Umgebungsschnittstelle. Mittels Sensoren kann ein aktuelles Realitätsmodell erzeugt werden und durch Aktoren können Veränderungen der Umwelt erreicht werden (Pervasive Computing) [Matt01]. Drittens bieten reaktive und proaktive Eigenschaften die Möglichkeit, dass Objekte selbstständig miteinander in Interaktion treten können (Autonomic Computing) [BaBe02].

Zusammengenommen ergibt sich für einzelne Objekte die Möglichkeit, ein stets aktuelles Abbild ihrer Umgebung zu erzeugen und diesem Abbild entsprechend ergeben sich umweltgerechte Handlungsoptionen. Objekte können autonom auf eine sich (ständig) verändernde Umwelt reagieren und sich selbst organisieren. Zentrale Informationssysteme, welche mit allen Daten versorgt werden müssen, können durch dezentrale, sich selbst organisierende Systeme abgelöst werden. Ein selbstorganisierendes System koordiniert sich selbst innerhalb eines gegebenen Regelrahmens, der die Aktionsmöglichkeiten der einzelnen Komponenten einschränkt. Die Beziehungen und Interaktionen innerhalb des Systems werden endogen hervorgerufen und ergeben systemweite Muster. Durch marktähnliche Verhandlungen werden Entscheidungen durch dezentrale Komponenten getroffen, denen nur unvollständige, direkt verfügbare Informationen sowie begrenzte Verarbeitungsmöglichkeiten zur Verfügung stehen. Eine Sicht auf das Gesamtsystem und ein Wissen über die Interdependenzen existiert nicht. Die Festlegung von Koordina-

tionsgrößen und des Regelrahmens sind die Grundlagen jeglicher Form von Selbstorganisation. [EySM03]

Prototypisch konnte die Machbarkeit sich selbst organisierender Prozesse auf Basis von UC-Technologien bereits gezeigt werden. Im Projekt EMIKA des Lehrstuhls Telematik der Universität Freiburg konnte in Experimenten gezeigt werden, dass die Terminplanung im Rahmen der Krankenhauslogistik durch eine dezentrale Selbstorganisation mit Hilfe von UC-Technologien und Software-Agenten realisiert werden kann [Eyma03; Muel03]. Weitere Modelle für sich selbst organisierenden Prozesse mittels UC-Technologien finden sich bspw. in [FrHR04; ZiWB06].

2.2 Charakteristiken und Nutzenpotenzial von sich selbst organisierenden Prozessen

Dem Idealbild eines sich selbst organisierenden Prozess liegen vier Charakteristiken zu Grunde: [Prob92; Prob87]

Autonomie: Sich selbst organisierende Systeme sind autonom, da die Gestaltung und Lenkung aus dem System selbst und nicht von außen geschieht. Eine Austauschbeziehung mit der Umwelt besteht aber selbstverständlich weiterhin.

Komplexität: Sich selbst organisierende Systeme sind komplex, da ihre einzelnen Teile durch wechselseitige, sich permanent ändernde Beziehungen miteinander vernetzt sind. Das globale Ordnungsmuster ist ein Resultat aus der Interaktion der einzelnen Teile. Die hohe Interaktivität und die der Selbstorganisation zugrunde liegenden Handlungsspielräume der einzelnen Teile führen zur Unmöglichkeit der vollständigen Beschreibbarkeit sowie Vorhersagbarkeit des Verhaltens der Systemen [Foer84a, Foer84b].

Selbstreferenz: Sich selbst organisierende Systeme sind selbstreferentiell, da jedes Verhalten des Systems auf sich selbst zurück wirkt und wieder Ausgangspunkt für weiteres Verhalten wird. Störungen bzw. Veränderungen wird aus dem System selbst heraus begegnet. Die Aktivitäten des Systems sind das Resultat der „inneren“ Zusammenhänge.

Redundanz: In selbstorganisierenden Systemen erfolgt keine Trennung zwischen organisierenden, gestaltenden oder lenkenden Teilen. Alle Teile des Systems stellen potentielle Gestalter dar. Ordnung ist das Resultat verteilter Aktivitäten. Mehrere Teile im System haben gleiche bzw. ähnliche Fähigkeiten. Informationen werden über das System verteilt aufgenommen und verarbeitet und Entscheidungen werden dezentral getroffen.

Die Nutzenpotenziale solcher sich selbst organisierenden Prozesse sind insbesondere:

Flexibilität und Dynamisierung: Durch kleine und schnelle Regelkreise sowie der Vermeidung langer Kommunikations- und Weisungswege kann eine Steigerung der Anpassungsfähigkeit und somit der Flexibilität des Systems erreicht werden. Des Weiteren wird eine Früherkennung von Umweltveränderungen sowie eine pro- und interaktive Umweltgestaltung gestärkt (Dynamisierung). [Prob92; Fres94]

Beherrschung von Komplexität: Prozesse werden in ihrer Dynamik und Struktur immer komplexer. Dadurch wird die Bereitstellung aller für die Koordination notwendigen Informationen auf zentraler Ebene oftmals unmöglich [HoLS06]. Ein sich selbst organisierender Prozess kann durch die dezentrale Nutzung der Informationen und der Vermeidung einer zentralen Instanz, welche sämtliche Informationen sammeln und auswerten muss, die ansteigende Komplexität betrieblicher Prozesse beherrschbar machen. [MüMW04]

Robustheit: Sich selbst organisierende Prozesse weisen eine hohe Robustheit gegenüber Störungen und Fehlern auf. Fällt beispielsweise ein Teil des Systems aus, ist das System auf Grund der Eigenschaften Selbstreferenz und Redundanz in der Lage sich selbständig Alternativen zu suchen. [MüMW04]

3 Problemstellung und Zielsetzung: *Steuerung der Selbstorganisation*

Mit einer Selbstorganisation von Unternehmensprozessen sind jedoch auch erhebliche ungelöste Probleme verbunden, die eine Umsetzung der bisherigen (prototypischen) Modelle in die betriebliche Realität bisher verhindern. Selbstorganisierende Koordinationsformen führen zu emergenten Strukturen¹. Selbstorganisation in Unternehmen ist aber nur denkbar, wenn sichergestellt werden kann, dass das Ergebnis als wünschenswerte Ausprägung im Hinblick auf die Unternehmensziele angesehen werden kann. Der Vision von sich selbst organisierenden Prozessen, die ohne zentrale Planung und Überwachung zu einem wünschenswerten Ergebnis gelangen, steht der Albtraum eines autonomen Systems gegenüber, welches unkontrolliert seinen eigenen „Willen“ durchsetzt [MuMW04].

Es ist nicht möglich das Verhalten eines sich selbst organisierenden Prozess vollständig zu beschreiben bzw. vorhersagen [Foer84]. Dieses Problem wird durch die hohe Interaktivität,

¹ Hayek würde dies als spontane Ordnung bezeichnen [Haye68].

Dynamik und Komplexität eines UC-Systems weiter verstärkt. Ein solcher Prozess ist in seinem Ergebnis und Rückwirkungen, für Unternehmen nicht ausreichend beherrschbar und stellt dann auch keine Option für den betrieblichen Einsatz dar. Die Autonomie von Entscheidungen führt zudem zur Unberücksichtigung von Entscheidungsinterdependenzen zwischen (autonomen) Einheiten. Entscheidungsinterdependenzen können Grundsätzlich auf sequentielle Verknüpfung oder auf Grund von Überschneidung von Entscheidungsfeldern zurückgeführt werden [Fres05]. *Sequentielle Verknüpfungen* sind dadurch gekennzeichnet, dass die Entscheidung einer Einheit - auf Grund innerbetrieblicher Leistungsverflechtungen – die interne Umwelt einer anderen Einheit zielrelevant verändert. Die Situation liegt z.B. vor, wenn die Entscheidung der Produktionseinheit über das Ausbringungsvolumen die Festlegung der Beschaffungsmenge durch die Beschaffungseinheit bestimmt. Die Beschaffungseinheit würde in Kenntnis einer Produktionsausweitung eine andere Beschaffungsentscheidung fällen als ohne Information über diese Veränderung. *Überschneidungen von Entscheidungsfeldern* sind dadurch gekennzeichnet, dass Entscheidungen einer Einheit zu Veränderungen des Entscheidungsfeldes einer anderen Einheit führen. Die trifft bspw. bei der gemeinsamen Nutzung knapper Ressourcen zu. Eine dezentrale Koordination von Prozessen bedeutet zwangsläufig eine Nichtberücksichtigung von Entscheidungsinterdependenzen zwischen Einheiten und führt zu Kosten auf Grund des Verzichtes auf eine theoretische Optimallösung, wie sie bei einer Berücksichtigung der Gesamtinterdependenzen möglich wäre. Diese Kosten werden in der Organisationstheorie Autonomiekosten genannt [Fres05].

Im Rahmen der Controllingliteratur wird als Antwort auf Selbstorganisation der verstärkte Einsatz von Mechanismen des Selbstcontrollings gefordert [Horv95]. Dieses Selbstcontrolling muss aber insbesondere aufgrund des Interdependenzenproblem von Mechanismen des gesamtzielorientierten Steuerung flankiert werden. So stellt [Kies94] fest, dass die Voraussetzungen der betrieblichen Selbstorganisation insbesondere durch Fremdorganisation geschaffen werden müssen. Es ist erforderlich, ein Steuerungssystem zu entwickeln, welches selbstorganisierenden Prozessen förderlich ist, aber auch in der Lage ist, das Prozessergebnis in Richtung der Gesamtunternehmungsziele zu beeinflussen. [Kies94; Horv06]

Ziel der Dissertation ist die Entwicklung und Evaluation eines Steuerungsmechanismus für sich selbst organisierende Prozesse. Dieser Mechanismus hat zwei grundlegenden Anforderungen zu erfüllen (vgl. hierzu auch [Krys95; Kies94]):

Berücksichtigung der gesamtszielorientierten Steuerung: Bei der Steuerung betrieblicher Aktivitäten ist immer eine ganzheitliche – im Sinne der Unternehmensstrategie – Ausrichtung zu verfolgen. Diese Koppelung ist notwendig, um die Performance des Gesamtunternehmens zu optimieren. Zu Berücksichtigen sind hier bspw. die Ausnutzung von Synergieeffekten oder Interdependenzen aufgrund von Marktüberschneidungen. Ein Prozess muss daher mit entsprechenden Steuerungsmechanismen vertikal an die Unternehmensstrategie angeschlossen werden. [HaWi99; Wall99; Frie93]

„Marktkonforme“ Steuerung: Die Konzeption eines Steuerungssystems muss mit der Struktur des zu koordinierenden Systems korrespondieren [Horv06; Schi01; Hein76]. Daher sind bei der Entwicklung eines Steuerungsmechanismus auch die Charakteristika der Selbstorganisation explizit zu berücksichtigen. Um die Vorteile sich selbst organisierenden Handelns weitgehend zu erhalten ist es notwendig, eine möglichst „marktkonforme“ Steuerung zu ermöglichen. Das bedeutet, dass eine Steuerung über die Rahmenbedingungen des sich selbst organisierenden Prozess zu erfolgen hat und nicht mit Hilfe von Verbote oder Gebote.

4 State of the Art-Analyse: *Steuerung betrieblicher Prozesse*

Für die Steuerung betrieblicher Aktivitäten stehen eine Reihe von Verfahren zur Verfügung (Controllingkonzepte). Diese sind allerdings nicht in der Lage die aufgestellten Anforderung hinreichen zu erfüllen.

Bisherige Verfahren zur Unterstützung der Bildung von Steuerungsgrößen können in übergreifende System und prozessorientierte System unterschieden werden.

Übergreifende Systeme unterstützen die Gesamtzielausrichtung und sind in der Regel im Rahmen einer zentralen (Fremd-) Steuerung implementiert. Die bedeutendsten Systeme sind: [Küpp97; Adam96]

Zentralistische Führungssysteme: Im Rahmen von zentralistischen Führungssystemen sind Entscheidungs- und Weisungsrechte stark zentralisiert. Entscheidungsfelder werden möglichst wenig zerlegt um so Interdependenzen berücksichtigen zu können. Die Planung ist durch eine Top-Down Ansatz und einem hohen Standardisierungsgrad gekennzeichnet. [Küpp97]

Budgetierung: Die Budgetierung ist ein Instrument der Planung mit dem die (zentral) erstellten Pläne vor allem in wertmäßige Pläne transformiert werden. Die Budgets stellen einen in wertmäßigen Größen formulierten Plan dar, der eine Entscheidungseinheit für eine bestimmte Zeitperiode mit einem bestimmten Verbindlichkeitsgrad vorgegeben wird. Je nach Art der Budgets wird die Autonomie von Einheiten eingeschränkt. [Horv06]

Kennzahlen- und Zielsysteme: Kennzahlen sollen in verdichteter, quantitativ messbaren Form über betriebswirtschaftlich relevante Sachverhalte informieren und gehören zu den klassischen Instrumenten des Controllings. Kennzahlen können eine Informations- und/oder Steuerungsaufgabe haben. Da einzelne Kennzahlen nur eine geringe Aussagekraft haben, werden in der Regel Kennzahlensysteme eingesetzt, in denen gegenseitig abhängige und einander ergänzende Kennzahlen zusammengefasst werden. Das wohl bekannteste System dabei ist das Du-Pont Systems. Ein weiteres und in den letzten Jahren sehr populär gewordenes Kennzahlen- und Zielsysteme ist die Methode der Balanced Scorecard. Im Rahmen der Balanced Scorecard werden Strategien, Unterziele und die kritischen Leistungsmaßstäbe in Form eines Top-Down-Prozesses ermittelt. Das Ziel ist die Entwicklung eines ausgewogenen Zielsystems. Deshalb werden in der Regel die finanziellen Ziele durch eine Kunden-, eine interne Prozess- und eine Lern- und Entwicklungsperspektive ergänzt. Die Anzahl der Perspektiven kann allerdings durchaus größer oder kleiner sein. Ergebniskennzahlen werden mit fortlaufenden Indikatoren durch Ursache-Wirkungshypothesen verknüpft. [KaNo97; Hoff99; Horv98]

Die übergreifenden Systeme bieten ein breites Spektrum für eine gesamtzielorientierte Steuerung. Im Rahmen von zentralistischen Führungssystemen sind Entscheidungs- und Weisungsrechte stark zentralisiert. Die Planung wird dadurch sehr komplex und ist in hohem Masse durch ein Top-Down Ansatz geprägt. Im Rahmen solcher Systeme ist eine Berücksichtigung der Autonomie von Einheiten nicht vorgesehen. Im Rahmen der Budgetierung erhalten Einheiten zwar mehr Entscheidungsspielräume innerhalb eines vorgegeben Ressourcenrahmens, eine explizite Berücksichtigung der Autonomie findet allerdings auch nicht statt. Kennzahlen- und Zielsysteme sind in der Regel auch durch einen Top-Down Ansatz geprägt. Eine Berücksichtigung der Autonomie findet auch hier nicht statt. So ist das Balanced Scorecard Konzepte durch eine Ausrichtung an das Top-Management geprägt, während ein herunter brechen auf untere Ebenen oft vernachlässigt wird. [Perl99]

Prozessorientierte Systeme können die dezentrale (Selbst-) Steuerung einzelner Prozesse unterstützen. Diese Systeme gehen dem Grundgedanken nach, dass das primäre Steuerungsziel nicht mehr die Erfüllung einer speziellen Funktion, sondern in der Durchführung von Prozessen gesehen werden muss. Die bedeutendsten Systeme sind: [Biel94; ScVr94]

Basissysteme: Die so genannten Basissysteme dienen zur Planung, Steuerung und Kontrolle einzelner Performanceparameter (i.d.R. Kosten, Zeit und Qualität). So schafft das Prozesskostenmanagement durch die systematische Ermittlung der Aktivitäten und Leistungen einzelner Einheiten eine Prozesstransparenz und bietet durch die Ermittlung von Kostentreibern und Prozesskostensätzen griffige Ansatzpunkte für ein Kostenmanagement. Im Rahmen des Time Based Managements geht es um die Planung, Steuerung und Kontrolle der zeitlichen Dimension von Prozessen. Hierbei steht insbesondere die Optimierung der Durchlaufzeiten im Vordergrund. Aufgabe des Qualitätsmanagements ist es das gewünschte Ausmaß an Übereinstimmung der Leistungserfüllung mit den zuvor aufgestellten Spezifikationen sicherzustellen. [Coop98; Stal98; Vogg99]

Integrative Instrumente: Um kunden- und marktorientierte Elemente sowie eine ständige Verbesserung der Steuerungsinstrumente zu implementieren existieren die so genannten integrativen Instrumente. So dient die Methode des Target Costing Prozesse auf die Kundewünsche auszurichten, indem die Kosten bestimmt werden, zu denen eine geplante Leistung mit spezifizierten Eigenschaften hergestellt werden muss, um bei erwarteten Absatzpreis ein angestrebtes Gewinnniveau zu erreichen. Dabei umfasst dieses Konzept auch Maßnahmen, welche die Erreichung gesetzter Ziele unterstützt. Die Kaizen Methode dient zu einer permanenten und schrittweisen Verbesserung von Produkten und Prozessen. [Horv06; Coop98; Klin99]

Die Prozessorientierten Systeme unterstützen zwar eine dezentrale Steuerung einzelner Prozesse, allerdings kann dabei nicht von einer marktkonformen Steuerung gesprochen werden, da die konkreten Steuerungsmaßnahmen bzw. Zielgrößen (bspw. geforderte Durchlaufzeit) zentral vorgegeben werden. Die Anforderungen an die Steuerung sich selbst organisierender Prozesse können somit auch durch solche Systeme nicht erfüllt werden. [Schi01; HoLa96; Camp89]

5 Inhaltliches und methodische Vorgehensweise: *Entwicklung und Evaluation eines Mechanismus zur Steuerung sich selbst organisierender Prozesse*

Wie aufgezeigt wurde, existiert bisher kein wissenschaftliches fundiertes Verfahren, welches den Anforderungen an eine Steuerung sich selbst organisierender Prozesse erfüllt. Diese Forschungslücke soll im Rahmen der Dissertation in drei Phasen geschlossen werden.

Phase 1 : Beschreibung eines Szenarios zur Selbstorganisation aus der Logistik

Aufgrund der Tendenz zu immer kleineren Bestellmengen je Artikel und einer immer häufiger stattfindenden Bestellung steigen die Anforderungen an Logistiksysteme stark an [Gude99]. Wesentliche Probleme einer zentral koordinierten Logistik sind Anpassungsfähigkeit und die Reaktionszeit bei Störungen im System. Dem steht die Vision der Selbstorganisation gegenüber. Im Rahmen der ersten Phase wird in Kooperationen mit Industriepartnern für die Transportlogistik ein realistisches Szenario entwickelt. Im Rahmen des Szenarios werden mit Hilfe von UC-Technologien (bspw. RFID) der Wareneingang, die Lagerung und der Warenversand alternativ beschrieben. Jede logistische Einheit (z.B. Päckchen, Paletten und Behälter) kann zu tragbaren Kosten mit UC-Technologien gekennzeichnet werden und enthält so neben Produktinformationen auch ein (individuelles) Transportziel. Somit wird ein Wechsel von einer zentralen Fremdsteuerung in hierarchischen Organisationsstrukturen hin zu einer Selbstorganisation in autonomen Subsystemen möglich. Wie Datenströme im Internet können logistische Einheiten dann selbstständig ihren Weg zum Ziel bei einer gegebenen Infrastruktur finden [HoLS06] [FrHR04]. Die Routenbestimmung und Ressourcenallokation findet durch marktähnliche Verhandlungen der logistischen Einheit an den Knotenpunkten (bspw. Distributionszentrum) statt. Die Entwicklung des hierfür notwendigen Koordinationsmechanismus ist Teil der ersten Phase. Theoretische Grundlage für diesen Koordinationsmechanismus bildet die - aus der Volkswirtschaftslehre bekannte - allgemeine Gleichgewichtstheorie. Aus dieser Theorie ist bekannt, dass jedes kompetitive, durch Verhandlungen und Tausch erzielte Marktgleichgewicht paretoeffizient ist (Erstes Wohlfahrtstheorem) [Vari03][Mank04]. Der für das Szenario zu entwickelnde Verhandlungsmechanismus soll einer walrasianischen Auktion entsprechen, da diese ein solches Gleichgewicht ermöglicht [Blau88].

Phase 2: Entwicklung des Steuerungsmechanismus

In der zweiten Phase wird ein geeigneter Steuerungsmechanismus für den sich selbst organisierenden Prozess entwickelt. Dabei wird davon ausgegangen, dass bisherige Ansätze von betrieblichen Steuerungsmodellen nicht übernommen werden können (vgl. Kapitel 4 sowie [Horv06; Horv95]). Der Steuerungsmechanismus soll eine gesamtszielorientierte Steuerung ermöglichen und somit insbesondere eine Berücksichtigung betriebliche Interdependenzen erlauben, gleichzeitig wird aber eine möglichst marktkonforme Lösung angestrebt, um die Vorteile einer Selbstorganisation nicht zu gefährden. Die Entwicklung des Mechanismus soll in zwei Schritten erfolgen:

a) Ermittlung des gesamtszielorientierten Steuerungsbedarf

Ziel der ersten Stufe ist es, den Steuerungsbedarf abzuleiten, indem die notwendige (strategierelevanten) Integration des betrachteten Prozesses bestimmt wird sowie Entscheidungsinterdependenzen aufgedeckt werden. Dafür sollen zunächst die aus Gesamtunternehmenssicht strategierelevanten Steuerungsdimensionen bestimmt werden. Die strategierelevanten Steuerungsdimensionen können aus den - in vielen Unternehmen vorhandenen - übergreifenden Controllingkonzepten (Vgl. Kapitel 4) abgeleitet werden. Sollte ein Unternehmen kein übergreifendes Controllingkonzept aufweisen, bietet sich die Übernahmen der Dimensionen der Balanced Scorecard an [Schi01]. Daran anschließend muss die Bedeutung der dieser Steuerungsdimensionen für den konkreten Prozess ermittelt werden und daraus der Steuerungsbedarf abgeleitet werden. Dafür kann auf kann auf Methodiken zur Bestimmung der notwendigen, strategierelevanten Integration von Steuerungsobjekten [Schi01] sowie auf Ansätze der Autonomiemessung und -bewertung [Fres05; Whis64; Klat70] aufgebaut werden.

b) Entwicklung des Mechanismus

In der zweiten Stufe wird der Mechanismus zur „marktkonformen“ Steuerung des sich selbst organisierenden Prozess entwickelt. Als theoretische Grundlage kann auch hierfür auf die allgemeine Gleichgewichtstheorie zurückgegriffen werden. Danach ist es möglich, dass jedes beliebige, pareto-effiziente Marktgleichgewicht erreicht werden kann (Zweites Wohlfahrtstheorem) [Vari03][Mank04]. Dafür ist es notwendig, vor Beginn der Marktaktivitäten (Verhandlungen) eine Steuerung der Anfangsausstattung vorzunehmen. Die Steuerung der Anfangsausstattung vor den Verhandlungen ist unabhängig von den Verhaltens-

selbst und hat damit auch keine verzerrende, marktschädigende Wirkung. Ein pareto-effizientes Verhandlungsergebnis bleibt gewährleistet. Andere Steuerungsmaßnahmen, wie bspw. eine Beeinflussung der Tauschpreise oder der Machtverhältnisse der Tauschpartner haben eine verzerrende Wirkung und sind somit nicht marktkonform [Vari03][Mank04]. Im Rahmen der Dissertation soll daher analysiert werden, wie die Anfangsausstattungen der Tauschpartner (im Szenario die logistischen Einheiten) ausgestaltet bzw. werden muss, um die in Stufe a) definierte gesamtzielorientierte Ausrichtung zu ermöglichen.

Phase 3: Evaluierung durch Simulation

In der dritten Phase soll mit Hilfe von Simulationen ermittelt werden, ob der in Phase 1 entwickelte Koordinationsmechanismus und der in Phase 2 entwickelt Steuerungsmechanismus eine Selbstorganisation des betrachteten Prozesses tatsächlich ermöglicht und ob eine Steuerung der Ergebnisse erreicht werden kann. Für die Simulationen wird die vom Lehrstuhl Telematik, Universität Freiburg, für die Projekte EMIKA, Avalanche und B2B-OS entwickelte Simulationsplattform genutzt und erweitert.

6 Zusammenfassung

Die Technologien des Ubiquitous Computing (UC) ermöglichen die Realisierung der Selbstorganisation von betrieblichen Prozessen. Mit einer betrieblichen Umsetzung sind allerdings noch erhebliche ungelöste Probleme verbunden. So ist es nicht möglich das Verhalten eines sich selbst organisierenden Prozess vollständigen zu beschreiben bzw. vorhersagen. Für Unternehmen ist dies problematisch, da dadurch eine Abstimmung mit weiteren innerbetrieblichen Prozessen oder eine Integration der Gesamtstrategie nicht möglich ist. Bisherige Instrumente der Steuerung betrieblicher Aktivitäten sind nicht in der Lage den Anforderungen der Steuerung eines sich selbst organisierenden Prozesses zu genügen. Ziel des Dissertationsvorhabens ist es deshalb, ein Steuerungsmechanismus zu entwickeln, welcher eine gesamtzielorientierte Ausrichtung der Prozesse erlauben. Der Mechanismus soll eine „marktkonforme“ Steuerung ermöglichen, um die Effizienz der Selbstorganisation nicht zu gefährden. Die Entwicklung des Steuerungsmechanismus ist in drei Phasen unterteilt. Im Rahmen der ersten Phase wird ein Szenario für eine auf UC-Technologie basierte Selbstorganisation der Transportlogistik beschrieben und

ein konkreter Koordinationsmechanismus entwickelt, welcher die Selbstorganisation logistischer Einheiten erlaubt. In der zweiten Phase erfolgt die Entwicklung eines „marktkonformen“ Steuerungsmechanismus für das betrachtete Szenario. In der abschließenden dritten Phase werden der Koordinationsmechanismus und der Steuerungsmechanismus mit Hilfe von Simulationen auf ihre Eignung für den betrieblichen Einsatz hin getestet.

Literaturverzeichnis

- [Adam96] Adam, R.: Planung und Entscheidung: Modelle – Ziele – Methoden. Wiesbaden 1996.
- [BaBe02] Banavar, G.; Bernstein, A.: Software Infrastructure and Design Challenges for Ubiquitous Computing Applications. In: Communications of the ACM 45/12 (2002), S. 92-96.
- [Biel94] Biel, A.: Controller's sich verändernde Umfeld – Mengendaten-Controlling. In: Controller Magazin 19 (1994) 4, S. 225-228.
- [Blau88] Blaug, M.: Walras' Gleichgewichtsmodell – Fortschritt oder Sackgasse. In: Recktenwald, H.; Engels, W.; Hax, H. u.a. (Hrsg.): Léon Walras' Lebenswerk – Eine kritische Analyse. Vademecum zu einem frühen Klassiker der ökonomischen Wissenschaft. Frankfurt a.M. 1998, S. 45-60.
- [Camp89] Camp, R.: Benchmarking: the search for industry best practices that lead to superior performance. Milwaukee 1989.
- [Coop98] Cooper, R.: Schlank zur Spitze: Mit Konfrontationsstrategien erfolgreich ins 21. Jahrhundert, München 1998.
- [Eyma03] Eymann, T.: Digitale Geschäftsagenten – Softwareagenten im Einsatz. Springer Xpert.press, Heidelberg 2003.
- [EySM03] Eymann, T.; Sackmann, S.; Müller, G.: Hayeks Katallaxie: Ein zukunftsweisendes Konzept für die Wirtschaftsinformatik?. In: Wirtschaftsinformatik 2/2003.
- [Fres94] Frese, E.: Zum Einfluß der „neuen“ Produktions- und Organisationskonzepte auf die Standortentscheidung international tätiger Unternehmen. In: Zülich, G. (Hrsg.): Vereinfachen und verkleinern – die neue Strategie der Produktion, Stuttgart 1994.
- [Fres05] Grundlagen der Organisation: Konzepte – Prinzipien – Strukturen. Wiesbaden 1998.
- [FrHR04] Freitag, M.; Herzog, O.; Scholz-Reiter, B.: Selbststeuerung logistischer Prozesse – Ein Paradigmenwechsel und seine Grenzen. In: Industrie Management, 20(2004)1, S. 23-27.
- [Frie93] Friedl, B.: Anforderungen des Profit-Center-Konzepts an Führungssysteme und Führungsinstrumente. In: Das Wirtschaftsstudium 22 (1993) 10, S. 830-842.

- [Foer84a] von Foerster, H.: Principles of Self-Organization – In a Socio-Managerial Context. In: Ulrich, H.; Probst, G. (Hrsg.): Self-Organization and Management of Social Systems, Heidelberg 1984.
- [Foer84b] von Foerster, H.: Erkenntnistheorie und Selbstorganisation. In: Delfin-Zeitschrift für Konstruktion, Analyse und Kritik, 4 (1984), S. 6-19.
- [Gude99] Gudehus, T.: Logistik – Grundlagen, Strategien, Anwendungen. Springer, Berlin et al. 1999.
- [HaWi99] Hax, A., Wilde, D.: The Delta Model: Adaptive Management for changing World. In: Sloan Management Review 40 (1999) 2, S 11-28.
- [Haye68] Hayek, F. A., Der Wettbewerb als Entdeckungsverfahren. In: Kieler Vorträge 56/1968.
- [Hein76] Heinen, E.: Grundfragen der entscheidungsorientierten Unternehmensführung. München 1976.
- [Hoff99] Hoffmann, O.: Performance Management. Bern 1999.
- [HoLa96] Horváth, P., Lamla, J.: Kaizen Costing. In: Kostenrechnungspraxis 40 (1996) 6, S. 225-340.
- [HoLS06] ten Hompel, M.; Libert, S.; Sondhof, U.: Dezentrale Steuerung für Materialflusssysteme am Beispiel von Stückgutförder- und sortieranlagen. In: Logistics Journal 2006.
- [Horv95] Horváth, P.: Selbstorganisation und Controlling. In: Führungskräfte und Führungserfolg – Neue Herausforderungen für das strategische Management. Gabler, Wiesbaden 1995, S. 255 – 267.
- [Horv98] Mit Balanced Scorecard Strategien erfolgreich umsetzen. In: Scheer, A.-W. (Hrsg.): Rechnungswesen und EDV, 19. Saarbrücker Arbeitstagung, Heidelberg 1998.
- [Horv06] Horváth, P.: Controlling. 10 Auflage, Verlag Franz Vahlen 2006.
- [KaNo97] Kaplan, R.; Norton, D.: Balanced Scorecard – Strategien erfolgreich umsetzen, Stuttgart 1997.
- [Kies94] Kieser, A.: Fremdorganisation, Selbstorganisation und evolutionäres Management. In: Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung 46 (1994) 3, S. 199-228.
- [Klat70] Klatzky, S.: Automation, Size, and the Focus of Decision-Making: The Cascade Effect. In: Journal of Business 43 (1970), S. 141-151.
- [Klin99] Klingebiel, N.: Performance Management – Performance Measurement. In: Zeitschrift für Planung 9 (1998) 1, S. 1-15.
- [Kryst95] Krystek, U.: Auf dem Weg zum Selbstcontrolling?. In: Gabler Magazin 8 (1995) 9, S. 26-30.
- [Küpp97] Küpper, H.-U.: Controlling-Konzepte, Stuttgart 1997.
- [LyYo02] Lyytinen, K.; You, Y.: Issues and challenges in ubiquitous computing: Introduction. In: Communications of the ACM 45/12 (2002), S. 63-65.
- [Mank04] Mankiw, N.: Grundzüge der Volkswirtschaftslehre, Stuttgart 2004.

- [Matt01] Mattern, F.: Ubiquitous Computing - Der Trend zur Informatisierung und Vernetzung aller Dinge. In: G. Rossbach (Hrsg.): Mobile Internet. Tagungsband 6. Deutscher Internet-Kongress, dpunkt Verlag, Heidelberg 2001, S. 107-119.
- [Muel06] Müller, G.: Privacy and Security in Highly Dynamic Systems. In: Communications of the ACM 49/ 9 (2006), S. 28-31.
- [Muel03] Müller, G.; Kreutzer, M.; Strasser, M.; Eymann, T.; Hohl, A.; Nopper, N.; Sackmann, S.; Coroama, V.: Geduldige Technologie für ungeduldige Patienten: Führt Ubiquitous Computing zu mehr Selbstbestimmung?. In: Mattern, F. (Hrsg.): Total Vernetzt, Springer, Heidelberg 2003.
- [MüMW04] Müller-Schloer, C.; Malsburg von der, C.; Würtz, R.: Organic Computing. In: Informatik Spektrum, 4/2004, S. 332-336.
- [Perl99] Perlitz, M.: Unternehmen durch „Performance Measuremet“ erfolgreich machen. In: Gabler Magazin 13 (1999) 2, S. 6-10.
- [Prob87] Probst, G.: Selbstorganisation und Entwicklung. In: Die Unternehmung 41 (1987) 4, S. 242-255.
- [Prob92] Probst, G.: Selbstorganisation. In: Frese, E. (Hrsg.) Handwörterbuch der Organisation. Pöschel, Stuttgart 1992, S. 2255-2269.
- [Schi01] Schindera, F.: e-Business und die Steuerung teilautonomer Organisationseinheiten, Wiesbaden 2001.
- [ScVr94] Scholz, R., Vrohling, A.: Prozeß-Redesign und kontinuierliche Prozessverbesserung. In: Gaitanides, M., Scholz, R., Vrohling, A., Raster, M. (Hrsg.): Prozeßmanagement: Konzepte, Umsetzungen und Erfahrungen des Reengineering, München 1994.
- [Stal88] Stalk, G.: Time – The Next Source of Competitive Advantage. In: Harvard Business Review 66 (1988) 4, S.41-51.
- [Vari03] Varian, H. R.: Intermediate microeconomics, New York 2003.
- [Vogg99] Voggenreiter, D.: Qualitätscontrolling in Unternehmen des Anlagengeschäfts, Frankfurt a.M. 1999.
- [Wall99] Waller, C.: Emerging Issues in Performance Measurement. In: Journal of Strategic Performance Measurement 3 (1999) 1, S. 37-40.
- [Weis91] Weiser, M.: The Computer for the Twenty-First Century. Scientific American, Vol. 265, No 3 1991, S. 94 – 104.
- [Whis64] Whisler, T.: Measurement Centralization of Control in Business Organizations. In: Copper, W., Leavitt, H., Shelly, M. (Hrsg.): New Perspective in Organization Research, New York 1964.
- [ZiWB06] Zimmermann, R.; Winkler, S.; Bodendorf, F.: Supply Chain Event Management With Software Agents. In: Kirn, S. et al. (Hrsg.): Multiagent Engineering – Theory and Applications in Enterprises. Springer, Heidelberg 2006, S. 155 – 175.

Geschäftliche Verträge als Quelle für Services in unternehmensübergreifenden Informationssystemen

Christian Brandt

Wirtschaftsinformatik 1 – Betriebswirtschaftliche Informationssysteme

Universität Paderborn

33098 Paderborn

Christian_Brandt@notes.upb.de

Abstract

Services werden aus geschäftlichen Verträgen abgeleitet, um bei Vertragsschluss Software zu erhalten, die unternehmensübergreifende Prozesse steuert. Es wird ein material-konkretes Datenmodell für Verträge, das auf juristischen Normen und technisch-betriebswirtschaftlichen Standards basiert, rekonstruiert; ein ergänzendes, formal-abstraktes bildet den Ausgangspunkt, Verträge aus materialen Bausteinen zusammenzusetzen. Mithilfe des material-konkreten Modells werden Services konzipiert.

1 Einleitung

Anforderungen neuer Geschäftsmodelle müssen schnell in Unternehmensprozesse umgesetzt werden, die durch Informationssysteme unterstützt werden [ScAE05]. Dies kann nur gelingen, wenn auf zeitaufwändiges „Programmieren im Kleinen“ verzichtet wird und die Systeme in einer übergeordneten Sprache („im Großen“) zusammengestellt werden [DeKr75]. Hersteller betriebswirtschaftlicher Informationssysteme schlagen u.a. serviceorientierte Architekturen (SOA) vor. Eine SOA ist ein abstraktes Software-Architektur-Konzept [DJMZ05, 8], das Anwendungen aus verteilten Eigen- und Fremdkomponenten (Services) zusammensetzt. Ein Service stellt Funktionalität unabhängig von seinen Implementierungsdetails bereit. Zur Laufzeit suchen Anwendungen situativ geeignete Services (lose Kopplung) und nutzen diese

über standardisierte Schnittstellen (dynamisches Binden). So sind Anwendungen in einer SOA funktional zerlegt, was eine prozessorientierte Betrachtungsweise erleichtert [DJMZ05, 8].

Eine maschinenlesbare, standardisierte Schnittstellenbeschreibung der Services auf syntaktischer Ebene reicht nicht aus, um geeignete zu finden und mit ihnen automatisiert zu interagieren. Eine übergreifende Semantik ist erforderlich. „Die Herausforderung im Umfeld der serviceorientierten Architekturen ist es daher, Semantik so eindeutig zu formulieren, dass menschliche Interpretationsvorgänge verzichtbar werden“ [DJMZ05, 290].

Unternehmensübergreifende Semantiken zu definieren ist keine originäre Aufgabe von IT. Im Geschäftsleben existieren Normen und Standards, die Beteiligten Fachsprachen anbieten. Juristische Normen werden seit Jahrhunderten entwickelt und angepasst. Sie regeln eine Vielzahl konkreter Geschäftsvorgänge durch Typisierung einzelner Geschehensabläufe und Konzentration auf das Wesentliche [Klun02, 13]. Nationale und internationale Standards vereinheitlichen Ressourcen, Produkte und Prozesse technisch und betriebswirtschaftlich unternehmens-, branchen- und industrieweit.

„Verträge beschreiben den Vorgang freier Austauschprozesse“ [Schw05, 180] zwischen Beteiligten. Sie bilden - basierend auf juristischen Normen und technisch-betriebswirtschaftlichen Standards - eine Brücke zwischen Sprachwelten der Vertragsparteien. „Verträge sind gewissermaßen die geborene ‚adaptive‘ Basis für die Geschäftsregeln eines Unternehmens, da ihre ‚Metasprache‘ die Sprache des Geschäftslebens ist“ [Fisc06, 141].

Das traditionelle Betrachten der physischen Güter- und Geldflüsse mithilfe von Instrumenten des Rechnungswesens und der Logistik lässt häufig nur einen statischen, z.T. vergangenheitsorientierten Blick auf die *physisch vorhandenen* Ressourcen, Produkte und Prozesse des Unternehmens zu. Das Betrachten der *vertraglich zugesicherten* Ressourcen, Produkte und Prozesse bietet einen Ausblick in die Zukunft und erlaubt, die Handlungsspielräume der Beteiligten näher zu bestimmen.

Verträge sind die häufigsten Rechtsgeschäfte zwischen mehreren Beteiligten [Klun02, 72] und unterliegen weitgehender Inhaltsfreiheit, so dass sie in der Praxis vielfältig gestaltet auftreten. Unter anderen untersuchen [ChCT03], [MSSW03] und [Xu04], wie Vertragsinhalte in Software zu verwandeln sind, um zur Laufzeit die Parteien bei der Vertragsüberwachung (contract monitoring), -erfüllung (contract fulfillment) oder -durchsetzung (contract enforcement) zu unterstützen. Sie modellieren Verträge allgemeingültig und formal, d.h. sie versuchen, große

Domänen („der Vertrag an sich“) „von oben“ semantisch zu greifen. Ihre Modelle haben noch nicht die Konkretisierungsstufe erreicht, um Verträge in der Realität umfassend abzubilden. Daher wirbt [Ortn99, 238] dafür, Software „konstruktivistisch‘ aus dem Handeln (Praxis), das zu einer zweckmäßigen ‚Versprachlichung‘ (Semiose) der Sachverhalte in den Anwendungsbereichen geführt hat, systematisch“ zu rekonstruieren („materialsprachliches Verfahren“ [Ortn02, 44]). Materiale Rekonstruktion bedeutet Konzentration auf eine spezifische Domäne (z.B. Kaufverträge einer Branche) und generiert viele spezifische Lösungen. Identische Teillösungen („Bausteine“) können mehrfach verwendet werden, wenn sie atomar, d.h. gesamtlösungsunabhängig, und auffindbar gespeichert werden. [MiBo95] schlagen vor, juristische Normen in einem Repository („legal rules repository“) zu speichern und passende Vertragsvorlagen („contract templates“) zu erstellen. In Anlehnung soll ein Repository aufgebaut werden, das es ermöglicht, neue Lösungen aus bestehenden Bausteinen zusammenzusetzen. Dazu muss ein „Bauplan“ existieren. Die Analyse der Bausteine kann helfen, Rückschlüsse auf formale Zusammenhänge zu ziehen und diese in einem formal-abstrakten Modell, das zum Ausgangspunkt einer Konfiguration mit „materialen Bausteinen“ wird, darzustellen. Eingesetzt werden sollen jedoch material-konkrete Modelle.

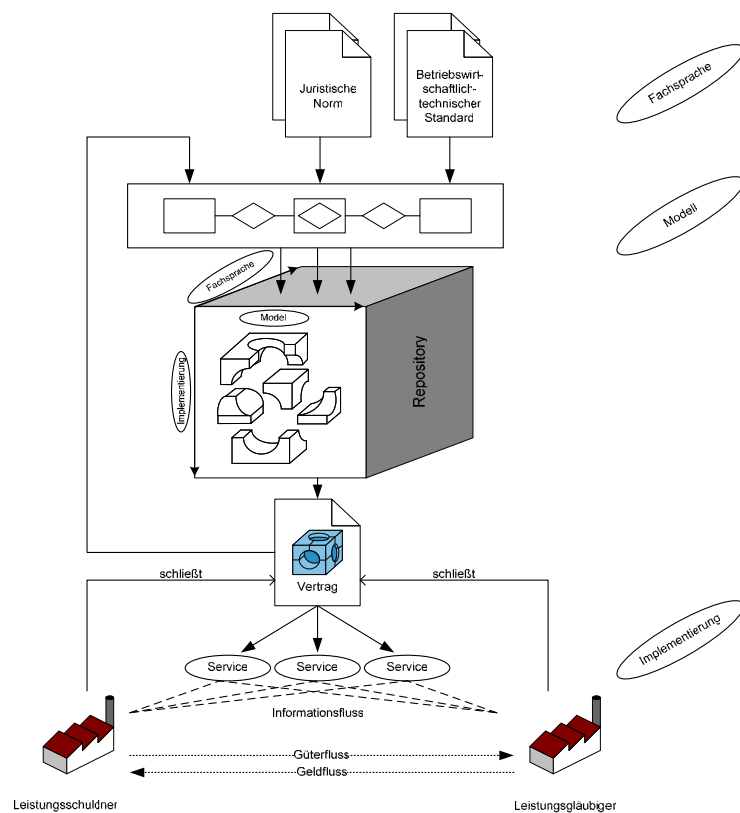


Abbildung 1: Architekturkonzept

Einige Bausteine schließen sich aus, andere bedingen einander. Konfigurationsmechanismen (z.B. aus [LeWi94]) müssen sicherstellen, dass nur gültige Verträge (contract validity [MiBo95]) gebildet werden können. Zusammengesetzte Bausteine bilden einen E-Contract und müssen in Services überführt werden. Sie steuern zur Laufzeit des Vertrages die Handlungen der Beteiligten und bewirken damit die Vertragserfüllung.

Im Folgenden wird ein materiales Datenmodell aus einem Beispielvertrag rekonstruiert. Parallel dazu wird aus zugrunde liegenden juristischen Normen und betriebswirtschaftlich-technischen Standards ein formales Modell entwickelt, dass zur Konfiguration dienen kann. Abschließend wird beispielhaft beschrieben, wie Services mithilfe des materialen Datenmodells die Handlungen der Vertragsparteien steuern.

2 Rekonstruieren des Datenmodells störungsfreier Vertragsabläufe

Die Rekonstruktion soll am Beispiel des störungsfreien Ablaufs, d.h. ohne die Betrachtung von Leistungsstörungen oder Vertragsbrüchen [Heus02, 292], eines Kaufvertrages (s. Abbildung 2) über normierte Messingbleche zwischen Unternehmen der stahl- und metallverarbeitenden Industrie stattfinden.

<p style="text-align: center;">Vertrag</p> <p>zwischen Fa. Deutsche Kupfer- und Messingwerke AG in Aachen, Heinestraße 21, vertreten durch ihre Vorstandsmitglieder Wilhelm Schmitz und Josef Müller, Verkäufer (im nachfolgenden kurz »Fabrik« genannt) und Fa. Josef Kraus GmbH, Feinmechanische Werkstätten in Düsseldorf, Neusser Straße 49, vertreten durch ihren Geschäftsführer Josef Kraus, Käufer (im nachfolgenden kurz »Firma« genannt).</p> <p style="text-align: center;">§1</p> <p>Die Fabrik verkauft der Firma zur Lieferung auf Abruf in der Zeit vom 1.7. bis 31.12.2007 10 Tonnen Messingblech CuZn37 (MS63) weichgebeizt, ziehfähig (R500), in fabrikationsgroßen Tafeln von etwa 2000x600x2mm, zum Preise von 255 € für 100 kg.</p> <p style="text-align: center;">§2</p> <p>Die Abrufmengen sollen auf die Lieferzeit gleichmäßig verteilt werden. Zur Lieferung von mehr als 2000 kg in einem Monat ist die Fabrik nicht verpflichtet.</p> <p style="text-align: center;">§3</p> <p>Die Bezahlung für jede Teillieferung ist mit 50% binnen 4 Wochen nach Lieferung, der Rest binnen 2 Monaten nach der ersten Zahlung zu leisten.</p> <p>[...]</p> <p style="text-align: center;">§10</p> <p>Erfüllungsort und Gerichtsstand ist Aachen. — Änderungen dieses Vertrages bedürfen der Schriftform.</p> <p>Aachen, 10.02.2007</p>

Abbildung 2: Beispielvertrag, ähnlich zu [NaSF80, 16]

Das materiale Datenmodell (s. Abbildung 3) beschreibt konkret das Verpflichtungs- und Erfüllungsgeschäft des Beispielvertrages. Das formale Datenmodell (s. Abbildung 4) bildet die

Basis, um materiale Datenmodelle für verschiedene Verträge zu konfigurieren. Beide zeigen ausschließlich unternehmensübergreifend relevante Objekte, Eigenschaften und Beziehungen, besitzen jedoch weit reichende Implikationen auf innerbetriebliche Ressourcen, Produkte und Prozesse.

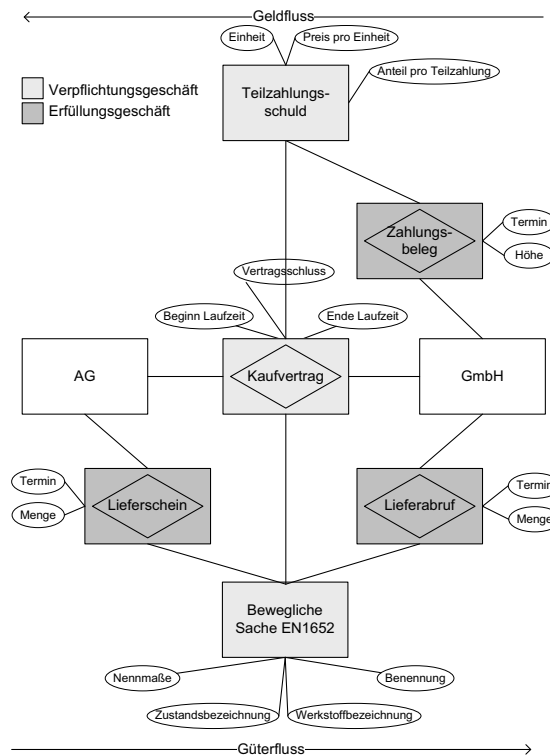


Abbildung 3: materiales Datenmodell des Beispielvertrages

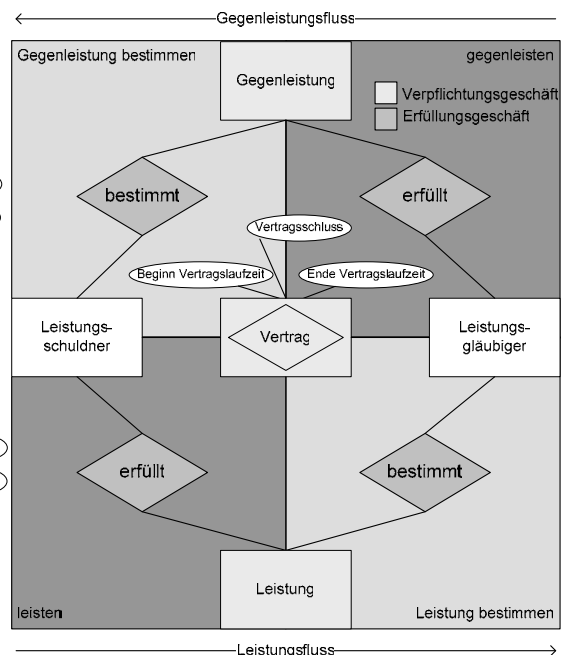


Abbildung 4: formales Datenmodell

Auf seiner horizontalen Achse zeigt das materiale Datenmodell eine AG (Käufer) und eine GmbH (Verkäufer), die durch den Kaufvertrag aneinander gebunden sind. Beide Rechtsformen verleihen den Vertragsparteien die Kaufmannseigenschaft. Formal-abstrakt repräsentieren die Entitäten Leistungsschuldner (gleichzeitig: Gegenleistungsgläubiger) und -gläubiger (gleichzeitig Gegenleistungsschuldner) die Vertragsparteien.

Juristische Normen qualifizieren Kaufverträge als Veräußerungsverträge und verpflichten den Verkäufer zum Übertragen von (beweglichen oder unbeweglichen) Sachen oder Rechten gegen Geld als Gegenleistung [Klun04, 311]. Auf der vertikalen Achse des materialen Datenmodells ist eine technisch-betriebswirtschaftlich standardisierte, *bewegliche Sache nach EN1652* [EN97] („10 Tonnen Messingblech CuZn37...“) durch den Kaufvertrag an eine *Teilzahlungsschuld* („255€ für 100 kg“) geknüpft. Sie zeigt das Verpflichtungsgeschäft. Der technisch-betriebswirtschaftliche Standard EN1652 [EN97, 6] legt die Attribute Nennmaße, Benennung,

Zustands- und Werkstoffbezeichnung als Bestellangaben für Kupfer- und Kupferlegierungen fest. Die Geldschuld ist mit Preis pro Einheit definiert.

Formal-abstrakt beschreibt ein Vertrag den Vorgang freier Austauschprozesse zwischen den Parteien [Schw05, 180] und begründet die Verpflichtung zur Leistung und Gegenleistung (vertikale Achse).

Im Erfüllungsgeschäft des Beispielkaufvertrages ruft der Käufer Lieferungen ab (Lieferabruf), der Verkäufer liefert (Lieferschein). Nach jeder Lieferung zahlt der Käufer (Zahlungsbeleg). Güter fließen vom Verkäufer zum Käufer, Geld in umgekehrter Richtung. Diese Rechtshandlungen bilden das Erfüllungsgeschäft und generieren Belege, die den Informationsfluss bilden. Sie sind in den Quadranten des materialen Datenmodells aufgeführt.

Formal-abstrakt erfüllen die Parteien ihre Verpflichtungen durch Rechtshandlungen, von denen je nach Vertrag entweder zwei, drei oder vier Typen existieren:

1. Der Leistungsgläubiger bestimmt die Leistung.
Sind Teile der Leistung bei Vertragsschluss offen gelassen (z.B. Termin, Menge, aber auch Inhalte), werden diese zur Laufzeit bestimmt [Klun04, 186]. Sonst entfällt diese Handlung.
2. Der Leistungsschuldner leistet
3. und bestimmt die Gegenleistung.
Ist die Gegenleistung bei Vertragsschluss nicht vollständig bestimmt, geschieht dies zur Laufzeit [Klun04, 187]. Sonst entfällt diese Handlung (s. Abbildung 3).
4. Der Leistungsgläubiger leistet gegen.

2. \ 1.		Leistung [Ware gegen...]		Gegenleistung [Geld gegen ...]	
		unbestimmt	bestimmt	unbestimmt	bestimmt
Leistung [Ware]	unbestimmt	[z.B. Tauschvertrag]		4 Typen	3 Typen
	bestimmt			3 Typen	2 Typen
Gg.leistung [Geld]	unbestimmt	4 Typen	3 Typen	[z.B. Kreditvertrag]	
	bestimmt	3 Typen	2 Typen		

Tabelle 1: Anzahl Typen von Rechtshandlungen nach Vertragsgestaltung

Das formale Datenmodell zeigt die Rechtshandlungen ebenfalls in den Quadranten. Das Erfüllungsgeschäft erstreckt sich von Beginn bis Ende der Vertragslaufzeit. Einige Verträge definieren einen Zeitpunkt, zu dem das Erfüllungsgeschäft durchgeführt sein oder werden muss (z.B. Fixhandelskauf §376 Abs. 1 HGB). Beginn und Ende der Laufzeit fallen dann zusammen.

Materialies und formales Datenmodell gegenüberzustellen zeigt, dass je nach Vertrag Leistung, Gegenleistung, Vertragsparteien und Belege aus ähnlichen Strukturen gebildet werden, aber juristisch unterschiedliche Konstrukte, betriebswirtschaftlich-technisch unterschiedliche Typen beschreiben.

Juristische Normen enthalten schwache Rechtswirkungssätze, die definitorisch und eigenschaftsbestimmend festlegen, wann „ein rechtlicher Zustand entsteht, erlangt wird, sich verändert oder beendet wird“ [Schw05, 48]. Sie schaffen juristische Konstrukte. Starke Rechtswirkungssätze beschreiben die Rechtsfolgen abhängig von rechtlichen Zustandskombinationen. Daher beeinflussen juristische Konstrukte den vertraglichen Ablauf. Beispielsweise erfordert ein Kaufvertrag über ein Grundstück (Konstrukt: unbewegliche Sache) notarielle Beurkundung. Die für den Vertragsablauf relevanten juristischen Konstrukte sind daher im Datenmodell abzubilden. Hierarchische Konstrukte können durch Generalisierung/Spezialisierung dargestellt werden; eigenschaftsbestimmende durch Attribute. Beispielsweise wird die Leistung eines Kaufvertrages als Kaufobjekt bezeichnet (s. Abbildung 5), das ein Recht oder eine Sache sein kann. Sachen lassen sich in bewegliche und unbewegliche Sachen (Grundstücke) unterscheiden. Bewegliche Sachen können die Eigenschaften Verbrauchbarkeit (§92 BGB) und Vertretbarkeit (§91 BGB) besitzen.

Betriebswirtschaftlich-technische Standards typisieren Gegenstände nach realen Sachverhalten. Sie werden in Verträgen referenziert und erlauben häufig erst die Anwendung schwacher Rechtswirkungssätze. Die durch diese Standards beschriebenen Eigenschaften sind ebenfalls in das Datenmodell aufzunehmen. Beispielsweise ist Messingblech CuZn37 eine bewegliche, verbrauch- und vertretbare Sachen, die unter den Typ Kupfer und Kupferlegierungen nach EN1652 [EN97] fallen.

Die in Abbildung 5 beispielhaft gezeigten Typen und Konstrukte sind gegen die jeweiligen Entitäten des formal-abstrakten Datenmodells auszutauschen, so dass ein materiales entsteht.

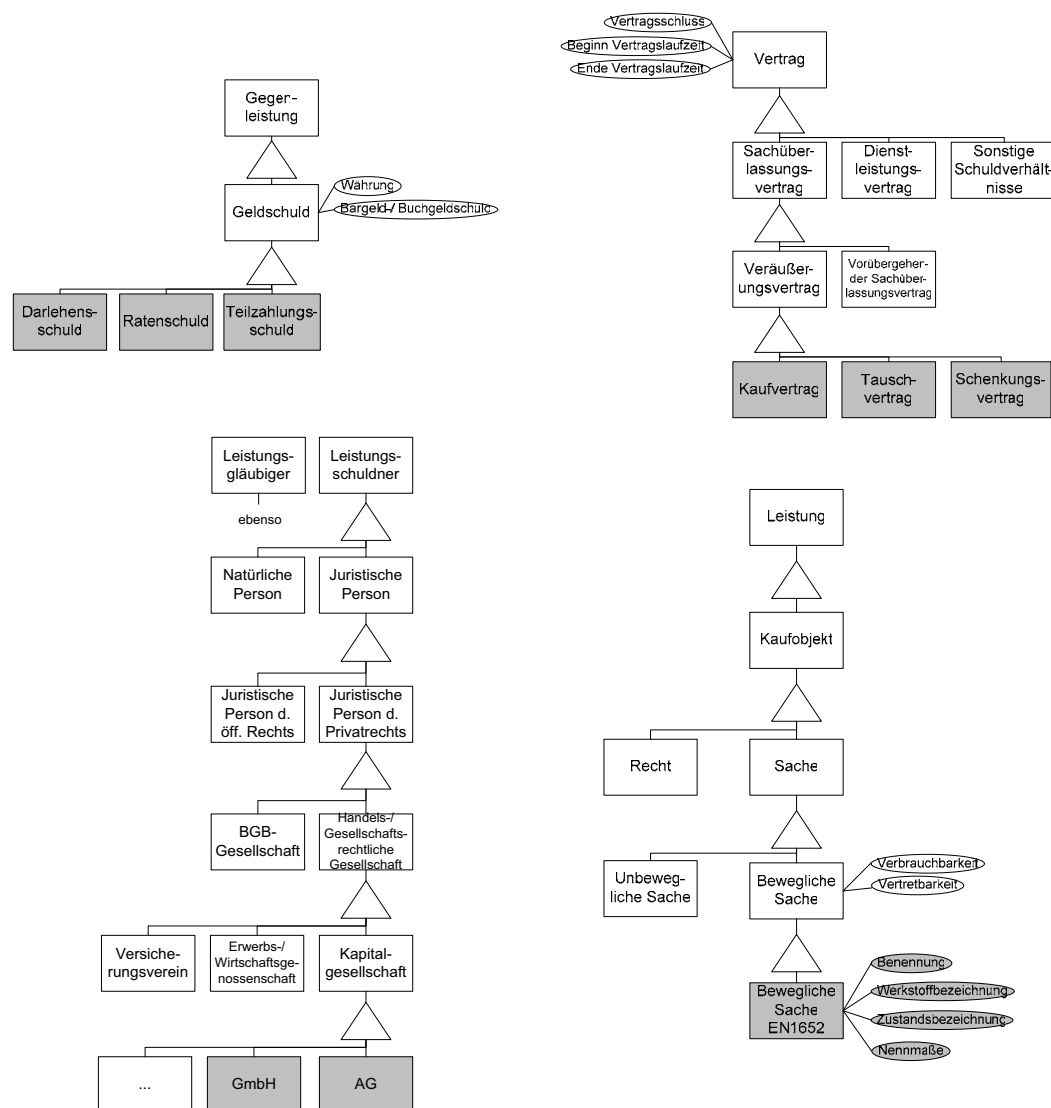


Abbildung 5: Konstrukte und Typen zum Bilden eines materialen Datenmodells

3 Modellieren unternehmensübergreifender Vertragsservices

Die in Abschnitt 2 genannten Typen von Rechtshandlungen sind durch die Vertragsparteien während des Erfüllungsgeschäftes auszuführen. *Leistung/ Gegenleistung bestimmen*, *Leistung/ Gegenleistung erfüllen* bilden zeit-, benutzer- und vertragsabhängige Services gesetzlicher Ebene, da sie im Allgemeinen Schuldrecht beschrieben sind. Darunter folgt die Vertragsebene mit Services, die aus dem Vertragsdokument hervorgehen (z.B. *Bewegliche Sache abrufen*). Im Vertrag nicht genau benannte, unbestimmte Inhalte müssen von den Vertragsparteien selbst

durch Ausführen von Services der Auslegungsebene (z.B. *Abrufmenge bestimmen*) bestimmt werden.

Der Leistungsgläubiger ruft den Service *Leistung bestimmen* mit Angabe einer Vertragsidentifikation auf. Auf Grundlage des materialen Vertragsdatenmodells wird ermittelt, ob der Nutzer den Service zu diesem Zeitpunkt ausführen darf (permission) oder es ihm verboten ist (prohibition). Vertragsparteien können auch verpflichtet sein (obligation), einen Service fristgerecht zu nutzen. *Leistung bestimmen* darf grundsätzlich während der Vertragslaufzeit ausgeführt werden. Der Service sucht in einem Repository selbständig nach geeigneten Services auf tieferen Ebenen. Aus dem Datenmodell geht hervor, dass der Vertrag die Ausprägung „Kaufvertrag“, die Leistung die Ausprägung „bewegliche Sache“ besitzt. Das Repository muss solche Ordnungsmerkmale aufweisen, so dass der Service *bewegliche Sache abrufen* ermittelt wird (Vertragsebene). Der Vorteil dieses Vorgehens ist, dass Servicenutzer keine detaillierten Kenntnisse über den Vertragsinhalt besitzen müssen. Durch Aufrufen der Services der Gesetzesebene werden tiefer liegende Services, die ausgeführt werden dürfen (permission) oder müssen (obligation), automatisch ermittelt.

Bewegliche Sache abrufen wird nach erfolgreicher Ausführung den Beleg „Lieferabruf“ produzieren, der die Attribute Abrufmenge und -termin besitzt. Darüber hinaus zu bestimmende Inhalte der Leistung könnten durch hinzufügen weiterer Lieferscheinattribute abgebildet werden. Die Services *Liefermenge/-termin bestimmen/annehmen* (Auslegungsebene) werden ermittelt. Der Leistungsgläubiger muss im vertraglich vereinbarten Handlungsrahmen die *Liefermenge bestimmen* und den *Liefertermin bestimmen*. Der Leistungsschuldner muss im vertraglich vereinbarten Handlungsrahmen die bestimmte *Liefermenge abwägen* und den bestimmten *Liefertermin abwägen*. Dabei kann der vereinbarte Handlungsrahmen im Vertrag nicht eindeutig spezifiziert sein. Dann ist dieser durch Leistungsgläubiger und -schuldner zu interpretieren. Unterschiedliche Interpretationen können zu Dissens führen und Störungen im Vertragsablauf hervorrufen. So kann z.B. der Leistungsgläubiger davon ausgehen, dass der -schuldner eine Lieferzeit von 3 Werktagen einhalten muss. Dieser geht jedoch von 5 Werktagen aus.

	Vertraglicher Handlungsrahmen	Interpretation Leistungsgläubiger	Interpretation Leistungsschuldner
Liefertermin	$\geq 01.07.2007$	dto.	dto.
	> heute + „übliche“ Lieferzeit in Werktagen	> heute + 3 Werktage	> heute + 5 Werktage
	$\leq 31.12.2007$ – „übliche“ Lieferzeit in Werktagen	$\leq 31.12.2007 - 3$ Werktage	$\leq 31.12.2007 - 5$ Werktage
Liefermenge	Liefermengen gleichmäßig verteilen	Liefermenge der ersten 5 Abrufe zwischen 1,5 und 1,9t	Standardabweichung Liefermenge <100kg
	Monatliche Liefermenge > 2000kg → Lieferung kann abgelehnt werden	Monatliche Liefermenge \leq 2000kg abrufen	Monatliche Liefermenge \leq 2050kg annehmen

Tabelle 2: Interpretation des vertraglichen Handlungsrahmens durch die Vertragsparteien

Da der Handlungsrahmen teilweise durch Interpretation unscharf ist [Fisc06, 139] und auch durch vertragsexterne Faktoren (z.B. langjährig bestehendes Geschäftsverhältnis, Nachfragermacht [Eger95, 236], Annahme über Interpretation der Gegenpartei etc.) beeinflusst wird, müssen Services der Auslegungsebene den Parteien die Möglichkeit geben, „weich“ zu reagieren. Dazu könnten die Parteien Toleranzbereiche für automatisiertes Annehmen, automatisiertes Ablehnen und Entscheidung durch einen Sachbearbeiter definieren.

Wurden Liefermenge und -termin bestimmt und angenommen, lösen die Services Fertigmeldungen aus. Auf diese reagiert der darüber liegende Service *bewegliche Sache abrufen*, erzeugt mit den gesammelten Daten einen Lieferabruf und löst eine Fertigmeldung aus. Abschließend meldet *Leistung bestimmen* die Fertigstellung. Dieses Ereignis ändert den Status des Service *Leistung erfüllen*. Dieser durfte zunächst nicht ausgeführt werden (prohibition); muss nun aber vom Leistungsschuldner zum Liefertermin genutzt werden (obligation).

Die Services *Leistung erfüllen*, *Gegenleistung bestimmen* und *Gegenleistung erfüllen* sind Gegenstand der weiteren Untersuchung.

4 Zusammenfassung und Ausblick

Abschnitt 2 erläutert, wie juristische Normen und betriebswirtschaftlich-technische Standards in ein material-konkretes Vertragsdatenmodell einfließen. Formale Zusammenhänge werden in

einem parallel entwickelten Datenmodell gezeigt. Beide Modelle sollen erweitert werden, um auch häufig vorkommende, betriebswirtschaftlich relevante Vertragsstörungen abzubilden. Materiales und formales Datenmodell gegenüberzustellen, bildet einen Ausgangspunkt, um über das Zusammensetzen von Vertragsservices aus materialen Bausteinen nachzudenken. Ziel der weiteren Untersuchungen ist es daher, ein Repository dieser Bausteine aufzubauen. Dabei werden einen Vertragstyp betreffende (z.B. Kaufvertrag) Bausteine betrachtet, die seine Leistung/ Gegenleistung oder seine Gestaltung ändern. Solche, die andere Vertragstypen (z.B. Mietvertrag) oder -parteien (z.B. Nichtkaufleute) berühren, werden nicht untersucht. Die Bausteine sind jeweils aus Sicht ihrer fachsprachlichen Bedeutung, aus Modellierungssicht und aus Implementierungssicht zu untersuchen. In Abschnitt 3 wurde ein Vorschlag präsentiert, wie Services auf Basis eines materialen Datenmodells ablaufen könnten. Dies liefert einen Ausgangspunkt für Untersuchungen der Implementierungssicht.

Literaturverzeichnis

- [ChCT03] Chiu, Dickson K.W.; Cheung, S.C.; Till, Sven: A Three-Layer Architecture for E-Contract Enforcement in an E-Service Environment. In: Proceedings of the 36th Hawaii International Conference on System Sciences (2003), S. 74ff
- [DeKr75] *DeRemer, Frank; Kron, Hans*: Programming-in-the-large versus programming-in-the-small. In: Proceedings of the International Conference on Reliable Software (1975), S. 114-121.
- [DJMZ05] *Dostal, W.; Jeckle, M.; Melzer, I.; Zengler, B.*: Service-orientierte Architekturen mit Web Services. Elsevier Spektrum Akademischer Verlag, München 2005.
- [Eger95] Eger, Thomas: Eine ökonomische Analyse von Langzeitverträgen. Metropolis-Verlag, Marburg 1995.
- [EN97] o.A.: EN1652:1997 – Kupfer und Kupferlegierungen – Platten, Bleche, Bänder, Streifen und Ronden zur allgemeinen Verwendung. Europäisches Komitee für Normung, Brüssel 1998.

- [Fisc06] *Fischer, Joachim*: Vertragsmanagement in Wertschöpfungsnetzwerken – Skizze eines Forschungsprogramms aus Sicht eines Wirtschaftsinformatikers. In: *Blecker, Thorsten; Gemünden, Hans-Georg (Hrsg.)*: Wertschöpfungsnetzwerke - Festschrift für Bernd Kaluza. Erich Schmidt Verlag, Berlin 2006, S. 129-146.
- [Heus02] Heussen, Benno: Vertragsmanagement. In: Heussen, Benno (Hrsg.): Handbuch Vertragsverhandlung und Vertragsmanagement. Verlag Dr. Otto Schmidt, Köln 2002, S. 27-310.
- [JeSc05] *Jertila, Aida; Schoop, Mareike*: The Language-Action Perspective and the Semantic Web - A Language Action Approach to Electronic Contracts. In: Proceedings of LAP (2005), S. 157-171.
- [Klun02] *Klunzinger, Eugen*: Einführung in das bürgerliche Recht. Vahlen, München 2002.
- [LeWi94] *Leist, Susanne; Winter, Robert*: Konfiguration von Versicherungsleistungen. In: Wirtschaftsinformatik 36 (1994) 1, S. 45-56.
- [MiBo95] *Milosevic, Zoran; Bond, Andy*: Electronic Commerce on the Internet: What is still missing? In: In Proceedings of the 5th Annual Conference of the Internet Society (1995), S. 245-254.
- [MSSW03] *Molina-Jimenez, Carlos; Shrivastava, Santosh; Solaiman, Ellis; Warne, John*: Contract representation for run-time monitoring and enforcement. In: IEEE International Conference on E-Commerce (2003), S. 103-110.
- [NaSF80] *Nath, Herbert; Schilling, Justin; Fingerhut, Michael*: Formularbuch für Rechtsgeschäfte – Verträge und Willenserklärungen insbesondere im kaufmännischen Bereich. Carl Heymanns Verlag, Köln 1980.
- [Ortn99] *Ortner, Erich*: Repository Systems. Teil 1: Mehrstufigkeit und Entwicklungsumgebung. In: Informatik-Spektrum 22 (1999) 4, S. 235-251.

- [Ortn02] *Ortner, Erich*: Sprachingenieurwesen – Empfehlung zur inhaltlichen Weiterentwicklung der (Wirtschafts-)Informatik. In: Informatik-Spektrum 25 (2002) 1, S. 39-51.
- [ScAE05] *Scheer, August-Wilhelm; Jost, Wolfgang; Wagner, Karl (Hrsg.)*: Von Prozessmodellen zu lauffähigen Anwendungen – ARIS in der Praxis. Springer, Berlin 2005.
- [Schw05] *Schwintowski, Hans-Peter*: Juristische Methodenlehre. Verlag Recht und Wirtschaft, Frankfurt am Main 2005.
- [Xu04] *Xu, Lai*: A Framework for E-markets: Monitoring Contract Fulfillment. In: *Bussler, Christoph; Fensel, Dieter; Orlowska, Maria E.; Yang, Jian (Hrsg.)*: Web Services, E-Business and the Semantic Web. Springer, Berlin 2003, S. 51-61.

Indikatoren für Einsatz und Evaluation von Multiagentensystem-Technologien zur effizienten Steuerung von Organisationsabläufen in Krankenhäusern

am Beispiel eines Prozesses in einem Krankenhaus in Deutschland und Indonesien

Dwi Anoraganingrum betreut von Prof. Dr. Torsten Eymann

LS Wirtschaftsinformatik (BWL VII)
Universität Bayreuth
95448 Bayreuth
dwi.anoraganingrum@uni-bayreuth.de

Kurzfassung

Das Krankenhaus als ein dynamisches System ist in viele einzelne teilautonome Funktionsbereiche untergliedert und bietet eine ideale Grundlage für die Anwendung von Multiagentensystemen (MAS). Mit deren Hilfe sollen Prozessabläufe effizienter gestaltet und die Qualität gesteigert werden. Zur Bestätigung dieser These, wird eine Prozessuntersuchung in einem Krankenhaus sowohl in Deutschland als auch in Indonesien zum Vergleich durchgeführt werden. Nach Untersuchung der Voraussetzungen, werden Indikatoren zur Effizienzmessung festgelegt, um ein Konzept zum Einsatz und zur Evaluation von MAS Technologien in einem Prozess eines Krankenhauses zu entwickeln. Dieser Artikel beschreibt Motivation, Stand der Forschung, Problemstellung, Zielsetzung, methodische Vorgehensweise und bisherige Arbeit.

1 Motivation

Jeder Mensch in jedem Land verdient es gesund zu sein, gerade deshalb ist das Gesundheitssystem eines Landes ein bedeutendes Element des sozialen Zusammenhalts und bildet heute in vielen Ländern die größte Dienstleistungsbranche. Im Hinblick auf die Konsequenzen dieser Systeme für *Effizienz, Effektivität und soziale Gerechtigkeit*, ihre Auswirkungen auf öffentliche Finanzen sowie ihre Fähigkeit, die Herausforderungen des medizinischen Fortschritts, der demographischen Entwicklung und der immer höheren Erwartungen zu bewältigen, sind kreative Politikansätze notwendig [OECD01].

Im Bereich Effizienz des Weltgesundheitsberichts (WHO) aus dem Jahr 2000 steht Deutschlands Gesundheitssystem auf Platz 25, Frankreich auf Platz 1, Italien auf Platz 2, USA auf Platz 37 und Indonesien auf Platz 92 von 191 Ländern, wobei im Bereich Gesundheitskosten Deutschland das drittteuerste Gesundheitssystem der Welt nach den USA und der Schweiz vorzuweisen hat [WHOR00; ETML01; JoWG01; TMLE01; FiVö03; DRGZ03; BRon04b; Refo04; Schr04; Suju04; NaBe05]. Länder mit den besten Gesundheitsniveaus entsprechen nicht immer den Ländern mit den effizientesten Gesundheitssystemen. Die Effizienzmessung bezieht sich auf Ausgaben für Gesundheit pro Kopf und ist insbesondere durch niedrige Ausgaben gekennzeichnet [WHOR00]. Die Methoden der Leistungsmessung mit Hilfe von IT-Lösungen liefern eine Grundlage politische Maßnahmen zu ergreifen, welche die Gesundheit verbessern. Reformen lassen sich durch sie kontrollieren [ETML01].

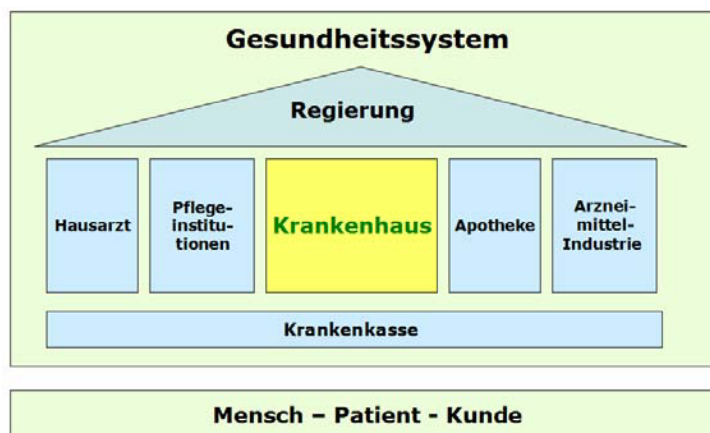


Abbildung 1 Die Elemente des Gesundheitssystems

Da das Krankenhaus, als ein Teil des Gesundheitssystems, seine Dienstleistungen direkt seinen Patienten anbietet, spielt es eine wichtige Rolle im Gesundheitswesen. Deshalb beeinflussen die Ausgaben der Krankenhäuser auch die Effizienz des Gesundheitssystems eines Landes.

Das Ergebnis einer Untersuchung von insgesamt 382 kommunalen, privaten Krankenhäusern sowie Universitätskliniken in Deutschland (01.10.2000 – 30.09.2005) von Peter Milde und Partner (PMP) hat durchgängig gezeigt, dass in über 62 Prozent der untersuchten Fälle zur falschen Zeit mit der falschen Personalmenge gearbeitet wird [Mild05]. Diese gravierenden Fehler in den Organisationsabläufen kostet insbesondere die öffentlichen Kliniken und Unikliniken in Deutschland Jahr für Jahr mindestens 30 Prozent mehr an Personalkosten als eigentlich erforderlich wären, von den Neben-, Sach- und Folgekosten und Wettbewerbsnachteilen ganz zu schweigen [BACR01; FiVö03; BRon04a; Refo04; Mild05].

Jedem Patienten sollte im Krankenhaus die optimale Behandlung zuteil werden [KTQO06]. Das Ziel von Kooperationen für Transparenz und Qualität im Gesundheitswesen (KTQ) kann erst erfolgen wenn alle Ressourcen, die für die Behandlung nötig sind, am richtigen Ort und zum richtigen Zeitpunkt zur Verfügung stehen.

In Deutschland sind im Gegensatz zu Indonesien die Ressourcen bereits vorhanden, sie werden allerdings noch nicht mit maximaler Effizienz genutzt [WHOR00; ETML01; JoWG01; TMLE01; FiVö03; DRGZ03; BRon04b; Refo04; Schr04; NaBe05]. Mit der Gesundheitsreform sowohl in Deutschland als auch in Indonesien werden gezielt neue und effizientere Versorgungsformen basierend auf den unterschiedlichen Ausgangspunkten der Länder gefördert [Grue96; WHOR00; JoWG01; Arga03; Suju03; Refo04; Suju04; NaBe05]. Es wird ein Wettbewerb um die beste Qualität in der Gesundheitsversorgung in Gang gebracht. Die Gesundheitsreform fordert und fördert, dass sich alle Partner im Gesundheitswesen einschließlich Krankenhäusern, Ärzten und Krankenkassen bewegen. Sie setzt Rahmenbedingungen, und sie schafft neue Chancen und Spielräume für eine nachhaltige Entwicklung der Gesundheitssysteme im jeweiligen Land [Refo04; NaBe05]. Allerdings sollten Effizienzverbesserungen arbeitsrechtlich realisierbar sein und die Qualität der Behandlung sollte nicht darunter leiden.

Reformmaßnahmen in Bezug auf Effizienz in Deutschland sind u. a.: Begrenzung der Verwaltungskosten der Krankenkassen, Korruptionsbekämpfung, Festbetragsregelung bei Arzneimitteln, freie Preise für nicht verschreibungspflichtige Arzneimittel und Honorarregelung für Apotheken. In der Qualitätssicherung setzt die Reform an sieben wichtigen Strukturveränderungen an: Hausarztssysteme, strukturierte Behandlungsprogramme, medizinische Versorgungszentren, Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, Qualitätsmanagement in Arztpraxen, Elektronische Gesundheitskarte und Integrierte Versorgung [Refo04].

In Indonesien leben etwa 240 Millionen Menschen auf einer Landfläche von ca. 1.912.988 km² [Mich06], allerdings mit einer regional sehr unterschiedlichen Bevölkerungsdichte. Auf diesem Hintergrund setzt die indonesische Gesundheitsreform an. Das Motto der indonesischen Gesundheitsreform lautet „Gesundes Indonesien 2010“ [Moel99; Suju03; Suju04; Mich06]. Dieses Programm beinhaltet: (1) saubere und gesunde Umwelt, gesünderer Lebensstil, Stärkung der Gemeinschaft, (2) die Vermittlung und Aufklärung über Gesundheitsprogramme, (3) Volksernährungsprogramme, (4) Personalentwicklung im Gesundheitssektor (5) Maßnahmen und Aufklärung gegen Missbrauch von Arzneimittel und Drogen (6) Verbesserung der Gesundheitspolitik und -management [Suju03; Suju04; Mich06].

Die Gesundheitsreform in Indonesien setzt dort an, wo in Deutschland bereits Ergebnisse erzielt wurden. Daher hat ein effizientes Gesundheitssystem in Indonesien eine andere Bedeutung. Mit etwa 240 Millionen Einwohnern hat Indonesien gerade einmal 120.000 Betten, die in 4 Klassen, A Allgemeine Krankenhäuser, 54 Klasse B, 213 Klasse C, 71 Klasse D, 335 private allgemeine Krankenhäuser, 1676 Puskesmas, 77 staatliche spezielle Krankenhäusern und 139 private Spezialkliniken eingeteilt werden [Moel99; WHOR00]. Gerade deshalb müssen Betten in Krankenhäusern effizient genutzt werden.

Nach den Gesetzesänderungen im deutschen Gesundheitswesen entstehen für Krankenhäuser bzw. Kliniken bei längerer Aufenthaltsdauer der Patienten höhere Kosten [DRGZ03; Refo04], da sich die Vergütung nur noch an festgelegten Pauschalpreisen je Behandlung orientiert. Daraus ergeben sich mehrere Konsequenzen wie beispielsweise: Schließung von Krankenhäusern, Fusion, Spezialisierung und höhere Effizienz.

Eine wichtige Rolle spielt hierbei der Operationsbereich, zum Beispiel Auslagerung ambulanter Operationen in spezielle günstigere ambulante OP-Räume. Eine andere Möglichkeit ist das verbesserte Zeit- und Personalmanagement mittels IT-Lösungen, den so genannten Multiagentensystemen.

2 Stand der Forschung

In diesem Abschnitt werden im Folgenden kurz einige Forschungen, die sich mit der Thematik Software Agent bzw. Multiagentensysteme Anwendungen im Gesundheitswesen beschäftigen, aufgeführt.

Software-Agenten als ein digitales Hilfsmittel, welches den Menschen direkt bei Entscheidungen unterstützt, sind autonom handelnde Softwareobjekte [Eyma03]. Der Begriff des Software-Agenten wird in Praxis und in der Forschung sehr vielfältig verwendet. Die Autonomie ist die grundlegende Eigenschaft, die einen Agenten von einem normalen Programm unterscheidet [Eyma03]. Autonomie wird als zielgerichtetes, proaktives und selbst-startendes Verhalten verstanden [Brad97].

Eigenschaften typischer Anwendungsdomäne von Multiagentensystemen sind wie in [Eyma03] beschrieben: verteilte Information, zentrales Sicherheitsrisiko, beschränkte Ressourcen, Integration existierender Systeme, schnelle Anpassung und hohe Skalierbarkeit.

Das Krankenhaus ist in viele einzelne (teil-)autonome Funktionsbereiche unterteilt, welche die Patienten – in Abhängigkeit ihrer jeweiligen Erkrankung – durchlaufen. Diese Funktionsbereiche sind entweder bestimmten Abteilungen organisatorisch direkt zugeordnet oder stehen funktional zentralisiert mehreren Abteilungen (gleichberechtigt) zur Verfügung, wie z.B. die Radiologie oder in zunehmenden Maße auch zentralisierte OP-Abteilungen [Schm99]. Diese funktionale Zentralisierung bietet zwar Möglichkeiten zur besseren Raum- und Gerätenutzung sowie einer höheren Flexibilität, jedoch führt diese dadurch bewirkte hohe Autonomie der einzelnen Funktionsbereiche insbesondere bei der gemeinsamen Nutzung eines zentralen OP-Komplexes zur Notwendigkeit umfassender Abstimmungen zwischen den einzelnen medizinischen Abteilungen [PHHH03].

Das Projekt MOKASSIN beschreibt die Anwendung der Multiagentensysteme in Prozessabläufen im Krankenhaus [KHJV98]. Die gemeinsamen Projekte aus dem DFG Schwerpunktprogramm „Intelligente Softwareagenten und betriebswirtschaftliche Anwendungs-Szenarien“ – SPP 1083, haben das Ziel interdisziplinäre Kooperationen zwischen Informatikern, Wirtschaftsinformatikern und Betriebswirtschaftlern in der Forschung so weit voranzutreiben, dass Agententechnologien für große Anwendungssysteme in realitätsnahen betriebswirtschaftlichen Anwendungs-Szenarien entwickelt und getestet werden können. Damit soll ein neuer, den Anforderungen der Dynamik weltweit verteilter Leistungsprozesse besser als bisherige Lösungen entsprechender Zugang zur Entwicklung von Anwendungssysteme sowie zu einer verbesserten Adaptivität von Unternehmen an dynamische Marktprozesse gefunden und erkundet werden [PHHH03].

Es gibt insgesamt sechs Projekte in SPP03. Das ASAINlog-Projekt greift das Problem der wachsenden Bedeutung informations-logistischer Fragestellungen auf. Das Projekt leistet wissenschaftliche Beiträge, durch die Fragen der Informationslogistik im Gesundheitswesen unter Nutzung von Multiagentensystemen beantwortet werden können. Ein essentielles Element des Lösungsmodells von ASAINlog ist die elektronische Patientenakte. Die Akte wird hierbei als „aktives Dokument“ bzw. Agent gesehen und nicht als Ressource. Es werden Dienste zur Behandlungsdokumentation von der Aufnahme bis zur Entlassung angeboten [Kirn02; PHHH03].

Das MedPAGE-Projekt befasst sich mit der funktionsbereichübergreifenden Planung, Steuerung und Koordination von Klinikprozessen (medical paths). Hierbei wird ein patientenzentrierter Ansatz gewählt, bei dem sowohl die Krankenhausressourcen als auch die Patienten als autonome Softwareagenten modelliert werden. Auf der Basis von Präferenzfunktionen verhandeln die Patientenagenten miteinander um die knappen Krankenhausressourcen. Dabei wird als Koordinationsmechanismus ein Marktmechanismus implementiert, in dem sich die Ressourcenallokation so lange verbessert, wie sich ein Agent durch eine Tauschhandlung besser stellen kann, ohne dass sich ein anderer Agent schlechter stellt. Durch realitätsnahe Tests und den Vergleich mit anderen Planungsmethoden wird die Nutzbarkeit der Agententechnologie untersucht [AwPa01; Kirn02; PJDH03; PHHH03].

Bei den Domänenmodellen des Projektes Agil2 steht die Notaufnahme im Mittelpunkt. Zentral ist hier ein Dienst zur Patientenaufnahme [Kirn02]. In diesem Projekt werden existierende Prozesse mit der Unterstützung eines javabasierenden Tools (genannt AGILShell) modelliert. Dieses Tool kann für das Design und die Implementierung von Multiagentensystemen benutzt werden. Der verfolgte Ansatz besteht aus drei Schritten: (1) Domänenexperten modellieren ihre existierenden Arbeitsprozesse; (2) Analyse der Prozesse, um Anwendungs-Szenarien für Agenten zu erkennen; (3) Optimierung der Prozesse durch Integration von Agenten. Basierend auf den existierenden Prozessen der einzelnen Projekte wird in interdisziplinärer Zusammenarbeit ein „agentifizierter“ Prozess erstellt, in dem Agenten vormals menschliche Aufgaben erfüllen. Hierbei verbessert die aktive Beteiligung des Benutzers am Entwicklungsprozess die Qualität der Software [BACR01; SJQF01; PHHH03].

Focus der ADAPT-Agenten sind Services zur Behandlungsplanung und zur verteilten Planung [Kirn02]. Das Projekt leistet einen Beitrag zur Lösung der dargestellten Studienplanungsproblematik. Primäres Ziel ist es, ein agentenbasiertes Simulationssystem zu erstellen, das in der Lage ist, die für die Durchführung klinischer Studien notwendigen Leistungsprozesse zu simulieren [HHPA03]. Zu diesem Zweck wurden entsprechende Simulationsmodelle implementiert, welche die notwendigen, am Prozess beteiligten Handlungsträger abbilden. Die Modelle basieren auf einer Akteur-zentrierten Sicht und wurden durch einen agentenorientierten Ansatz realisiert. Der entwickelte Prototyp unterstützt Mediziner und Mitarbeiter aus den an der klinischen Studie beteiligten Fachabteilungen bei der Analyse, Bewertung und Planung der Studienprozesse [HeHP00; PKHK00; HeHK02; Kirn02; PHHH03].

Das Projekt Policy-Agents leistet einen Beitrag zur Lösung der OP Terminplanungsproblematik. Durch den Einsatz eines agentenbasierten Planungssystems kann die OP-Terminplanung weitgehend automatisiert werden. Ein besonderer Fokus der Projekt Policy-Agents liegt dabei auf der expliziten Berücksichtigung der Interessen der einzelnen Abteilungen und beteiligten Personen. Dazu wird jeder Akteur durch seinen eigenen Softwareagenten repräsentiert, in dem ein Präferenzprofil des Akteurs hinterlegt ist. Im Rahmen der Terminplanung verhandeln diese Softwareagenten selbstständig miteinander und versuchen, eine effiziente Ressourcenallokation bei deutlich geringeren Transaktionskosten zu erzielen [PHHH03].

Der Fokus des Projekts Emika liegt auf der Echtzeitkoordination der Patientenlogistik in der Radiologie, so dass plötzlich auftretende Notfälle, wie hier angenommen, möglichst ohne Verzögerung in den bestehenden Zeitplan übernommen und zeitgenau untersucht werden können [PHHH03]. Emika bietet ein Benachrichtigungsservice, welcher Nachrichten für Akteure entgegen nimmt und über PDA oder Handy an die entsprechenden Empfänger ausliefert. Eine weitere Funktion ist die kontaktlose Lokalisierung/Tracking von Personen, welche über die Ausstattung der realen Akteure mit RFID-Chips (Radio Frequency Identification) realisiert wird [KHPP00; Kirn02; PHHH03; SaEM03; NiEy06]. Die Architektur des EMIKA-Systems ist dreistufig. Die höheren Schichten arbeiten mit den aggregierten Daten aus den unteren Schichten. So verringert sich die Datenmengen in jeder Schicht, deren Informationsgehalt steigt jedoch [NiEy06].

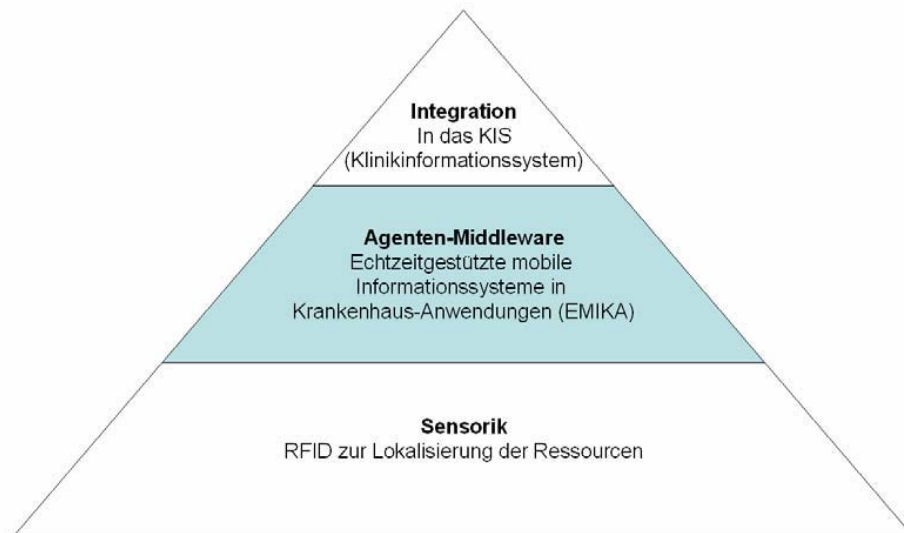


Abbildung 2 Die Architektur des EMIKA-Systems [NiEy06]

3 Problemstellung

Die Hypothese der Arbeit ist, dass die Anwendung von Multiagentensystemen zu einer Effizienzsteigerung bei der Steuerung eines ausgesuchten Prozesses in einem Krankenhausbereich beiträgt. Um diese Hypothese zu bestätigen, wird eine Untersuchung in dem ausgewählten Krankenhaus sowohl in Indonesien als auch in Deutschland durchgeführt.

Aus dieser Hypothese lassen sich folgende Fragen ableiten: Was versteht man unter Effizienz im Krankenhaus? Wie erreicht man Effizienz im Krankenhaus? Inwieweit ist die berechnete Effizienz arbeitsrechtlich realisierbar? Welche Vorteile und Nachteile bringt eine größere Effizienz für Krankenhaus und Patient sowohl in Deutschland als auch in Indonesien? Welche Indikatoren muss man betrachten, um diese Effizienz zu bewerten? Sind die Indikatoren in allen Ländern gleich? Können die Multiagentensysteme bzw. dezentralisierten Systeme einen Beitrag zu dem Ziel Effizienz in Krankenhäusern leisten? Welche Voraussetzungen müssen erfüllt werden, um MAS-Technologien zu realisieren? Effizienz als ein Teil der Qualitätssicherung führt zu bereits vorhandenen Zertifizierungen, wie z.B. KTQ, EFQM, ISO. Was bringt diese Zertifizierung für Krankenhäuser in Deutschland und in Indonesien?

4 Zielsetzung der Arbeit

Die Arbeit hat als Zielsetzung Indikatoren bzw. sog. KPI (Key Performance Indicator) festzulegen, die eine Messung und Evaluation von Multiagentensystemen und deren Einsatz ermöglichen. Ein umfassendes Konzept-Design für die Implementierung von MAS-Technologien wird anhand festgelegter Indikatoren entwickelt.

Dieses Ziel soll durch eine Untersuchung an ausgewählten Prozessen in einem Krankenhausbereich sowohl in Indonesien als auch in Deutschland erfolgen.

5 Methodische Vorgehensweise

Zunächst werden alle erforderlichen Grundlagen sowie Voraussetzungen für die Implementierung der MAS-Technologien in einem Krankenhaus mittels Literaturrecherche in der Voruntersuchungsphase erforscht.

Danach wird die Problemstellung der Arbeit auf einen Krankenhausbereich, der möglicherweise von MAS am stärksten profitieren wird, eingeschränkt. Die unterschiedliche Definition und Interpretation der Indikatoren für die Effizienz-Messung werden untersucht. Dies erfolgt durch die Literaturrecherche und die Begutachtung am jeweiligen Ort in Indonesien und Deutschland.

Im nächsten Schritt werden alle Anforderungen an die gewählte Multiagentensystem-Applikation sowie dessen Modellbeschreibung anhand der eingeführten Definitionen herausgearbeitet. Die Analyse bestehender Konzepte zur Modellierung und Simulation eines Szenarios und zu Prinzipien des Software-Designs wird durchgeführt. Die gefundenen Ansätze werden hinsichtlich der gestellten Anforderungen bewertet. Alle identifizierten Fehlbeträge werden zusammengefasst und als eigentliche Zielsetzung der Dissertation inklusive Handlungsempfehlung abgeleitet.

Die Konzeption der Modellbeschreibung erfolgt in einem weiteren Schritt. Dabei wird nach einigen grundlegenden Festlegungen zunächst eine Modellbeschreibung entwickelt, die den gestellten Anforderungen von festgelegten Indikatoren genügt und die durch das zu entwickelnde Konzept angewandt wird. Ein Ausblick weist auf weiterführende Folgearbeiten hin, die sich in dem Themenbereich anbieten.

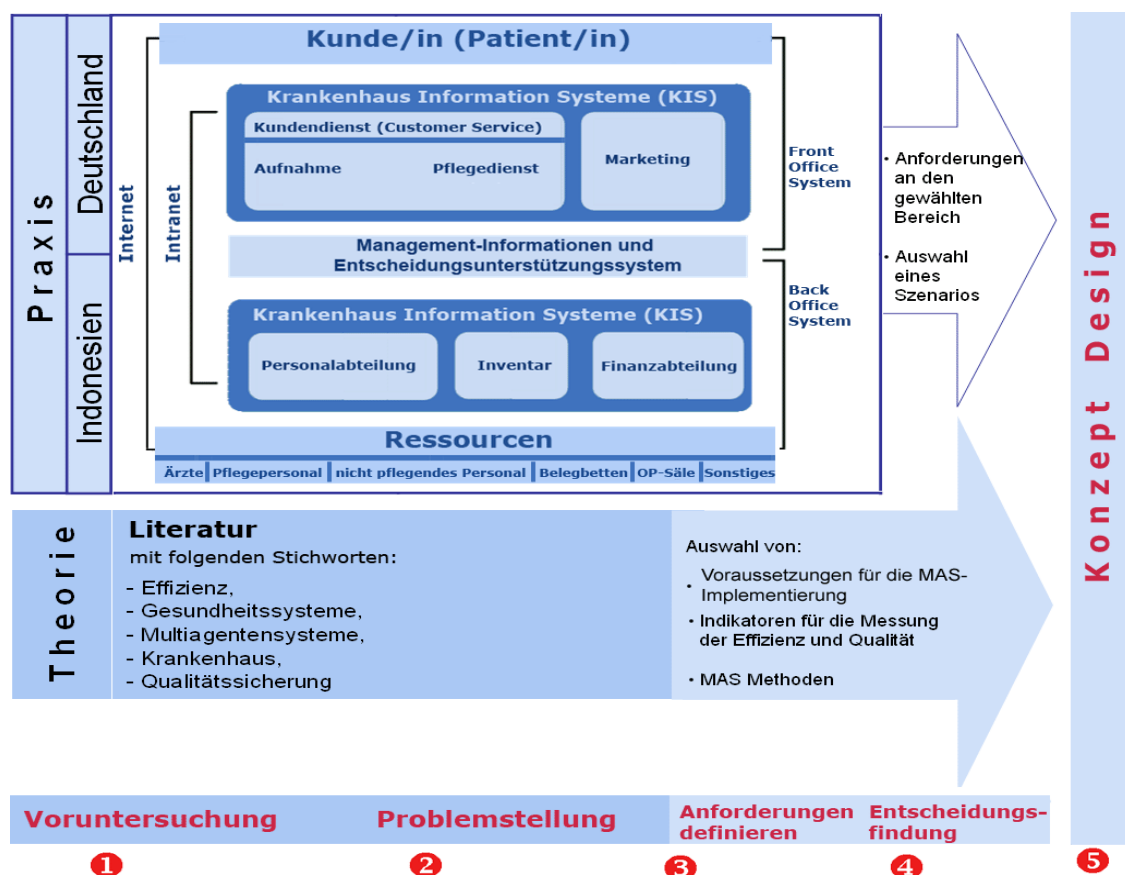


Abbildung 3 Methodischer Vorgehensweise

6 Bisherige Arbeit

Die erste Voruntersuchung wurde bereits in einem Krankenhaus in Indonesien durchgeführt [Anor06]. Es wurde festgestellt, dass vor der Implementierung der MAS-Technologien weitere Voraussetzungen erfüllt sein müssen. Hierzu gehören die Bereitschaft der Akteure, die Eignung der Infrastruktur und die Durchführung der Sensorik wie in Abbildung 2 dargestellt. Bei der Voruntersuchung wurden die OP-Bereiche (ICU und COT) als die Domänen identifiziert, die von dem Einsatz der MAS Applikation am stärksten profitieren könnten. Der nächste Schritt der Arbeit ist die Voruntersuchung in einem Krankenhaus in Deutschland, um dort den Ist-Zustand aufzunehmen und ein Soll-Konzept zu entwerfen.

Literaturverzeichnis

- [Anor06] *Anoraganingrum, Dwi*: Preliminary Study of Mas Technology Application in Wahidin Sudirohusodo Hospital. Jakarta, Bayreuth: Epos, 2006.
- [Arga03] *Argadiredja, Dadi S.*: Kebijakan Dan Strategi Desentralisasi Bidang Kesehatan. Jakarta: Departemen Kesehatan RI, 2003.
- [AwPa01] *Awizen, M.; Paulussen, Torsten*: Modellierung Von Kommunikationsprotokollen Für Die Dezentrale, Agentenunterstützte Koordination Von Krankenhausprozessen. In: (Hrsg): Informatik 2001: Wirtschaft und Wissenschaft in der Network Economie - Visionen und Wirklichkeit. Wien: Österr. Computer-Ges 2001, S. 883-88.
- [BACR01] *Beuscart-Zéphir, M.; Anceaux, F.; Cinquette, V.; Renard, J.*: Integrating Users' Activity Modeling in the Design and Assessment of Hospital Electronic Patient Records: The Example of Anesthesia. International Journal of Medical Informatics (2001), S. 151-71.
- [Brad97] *Bradshaw, Jeffrey M.*: Software Agents. Menlo Park, CA: MIT Press, 1997.

- [ETML01] *Evans, David B.; Tandon, Ajay; Murray, Christopher JL; Lauer, Jeremy A:* Comparative Efficiency of National Health Systems: Cross National Econometric Analysis. *BMJ* 323(2001), S. 307-10.
- [Eyma03] *Eymann, Torsten:* Digitale Geschäftsagenten. Berlin Heidelberg New York: Springer-Verlag, 2003.
- [FiVö03] *Fischer, Gerd; Völker, Lars:* Das Krankenhaus 2003: Prozessorientiertes Und Integriertes Dienstleistungsunternehmen? Rüsselsheim: EDS Consulting, 2003.
- [Grue96] *Grüneberg, Jürgen:* Sicherheit Für Mensch Und Umwelt. 75 Jahre Vde Bergisch Land. Verein zur Förderung von Innovation und Technologietransfer e.V., 1996.
- [HHPA03] *Heine, C.; Herrler, R.; Petsch, M.; Anhalt, C.:* Adapt - Adaptive Multi Agent Process Planning & Coordination of Clinical Trials. In: *AMCIS*. Tampa/ Florida 2003.
- [HeHK02] *Herrler, R.; Heine, C.; Klügl, F.:* Appointment Scheduling among Agents: A Case Study in Designing Suitable Interaction Protocols. In: *AMCIS*. Dallas 2002.
- [HeHP00] *Herrler, Rainer; Heine, Christian; Puppe, Frank:* Konzept Einer Schedulingkomponente Für Ein Krankenhaus-Multiagentensystem, 2000.
- [JoWG01] *John, J.; Wismar, M.; Geraedts, M.:* Aktuelle Forschungsperspektiven Von Gesundheitssystemforschung Und Gesundheitsökonomie in Deutschland. *Gesundheitswesen* 63(2001), S. 73-78.
- [Kirn02] *Kirn, Stefan:* Kooperierende Intelligente Softwareagenten. In: *Wirtschaftsinformatik* 44(2002), S. 53-63.
- [KHPP00] *Kirn, Stefan; Heine, Christian; Petsch, Mathias; Puppe, Frank; Klügl, Franziska; Herrler, Rainer:* Agentenorientierte Modellierung Vernetzter Logistikkreisläufe Als Ausgangspunkt Agentenbasierter Simulation - Beispiel Wöchnerinnenstation, 2000.

- [KHJV98] *Klauck, Christoph; Herzog, Otthein; Joeris, Gregor; Visser, Ubbo; Wache, Holger*: Mokassin Am Beispiel Der Anwendung Von Workflow-Technologie Im Krankenhaus. In: 1. Symposium Bremer Kliniken: Medizinische Forschung und Gesundheitswissenschaften. Bremen 1998.
- [KTQO06] *KTQ*: Transparenz Und Qualität Im Gesundheitswesen. http://www.ktq.de/ktq_ueber_uns/index.php, Abruf am 10.05.2006.
- [Mich06] *Michaud, Catherine M.*: External Resources Flows to the Health Sector in Indonesia. Jakarta: WHO, 2006.
- [Mild05] *Milde, Peter*: Kliniken Müssen Sich Öffnen Und Ihre Organisation Überdenken. <http://www.pmp-sc.com/modules.php?op=modload&name=News&file=article&sid=28>, Abruf am 10.05.2006.
- [Moel99] *Moeloek, F.A.*: Health Development Plan Towards Healthy Indonesia 2010. Jakarta: Ministry of Health, 1999.
- [NaBe05] *Nagel, Eckhard; Beske, Fritz*: Das Gesundheitswesen in Deutschland. Struktur – Leistung – Weiterentwicklung. Köln: Deutscher Ärzte-Verlag, 2005.
- [NiEy06] *Niemann, Christoph; Eymann, Torsten*: Softwareagenten in Der Krankenhauslogistik - Ein Ansatz Zur Effizienten Ressourcenallokation. eHealth 251(2006), S. 77-87.
- [DRGZ03] *NN*: Schnell Operiert, Schnell Entlassen. DRG-Zeitung 17.10.2003 2003.
- [Refo04] *NN*: Die Gesundheitsreform 2004. <http://www.die-gesundheitsreform.de>, Abruf am 10.05.2006.
- [BRon04a] *NN*: Leben 2020: So Viel Kostet Unsere Gesundheit. <http://www.br-online.de/leben2020/thema/gesundheitssysteme/gesundheitskosten.xml>, 06.12.2004, Abruf am 2006-05-10.

- [BRon04b] *NN: Leben 2020: Deutschland Im Internationalen Vergleich.* <http://www.br-online.de/leben2020/thema/gesundheitsysteme/internationaler-vergleich.xml>, 2004-12-06, Abruf am 2006-05-10.
- [OECD01] *OECD: Measuring Up: Improving Health Systems Performance in Oecd-Countries.* In: Smith, Peter (Hrsg). Ottawa, Canada: OECD 2001.
- [PJDH03] *Paulussen, Torsten; Jennings, N. R.; Decker, K. S.; Heinzl, A.: Distributed Patient Scheduling in Hospitals.* In: (Hrsg): 18th Int. Joint Conf. on AI. Acapulco, Mexico 2003.
- [PHHH03] *Paulussen, Torsten O.; Herrler, Rainer; Hoffmann, Astrid; Heine, Christian; Becker, Marc; Franck, Martin; Reinke, Thomas; Strasser, Moritz: Intelligente Softwareagenten Und Betriebswirtschaftliche Anwendungsszenarien Im Gesundheitswesen,* 2003.
- [PKHK00] *Puppe, Frank; Klügl, Franziska; Herrler, Rainer; Kirn, Stefan; Heine, Christian: Konzeption Einer Flexiblen Agentenkomponente Für Schedulingaufgaben Im Krankenhausumfeld,* 2000.
- [SaEM03] *Sackmann, S.; Eymann, Torsten; Müller, Günter: Emika - Real Time Controlled Mobile Information System in Health Care Applications,* 2003.
- [Schm99] *Schmidt-Rettig, B.: Profitcenter Und Prozessorientierung: Optimierung Von Budget, Arbeitsprozessen Und Qualität. Profitcenter-Organisation Und Prozessorganisation - Konflikt Order Konsens? Stuttgart,* 1999.
- [Schr04] *Schreyögg, J.: Gerechtigkeit in Gesundheitssystemen Aus Ökonomischer Perspektive. Gesundheitswesen,* 2004.
- [SJQF01] *Staccini, P.; Joubert, M.; Quaranta, J.; Fieschi, D.; Fieschi, M.: Modelling Health Care Processes for Eliciting User Requirements: A Way to Link a Quality Paradigm and Clini-Cal Information System Design. International Journal of Medical Informatics, (2001), S. 129-42.*

- [Suju03] *Sujudi, Achmad*: Indikator Indonesia Sehat 2010 Dan Pedoman Penetapan Indikator Provinsi Sehat Dan Kabupaten/ Kota Sehat. Jakarta: Departemen Kesehatan, 2003.
- [Suju04] *Sujudi, Achmad*: Sistem Kesehatan Nasional. Jakarta: Departemen Kesehatan RI, 2004.
- [TMLE01] *Tandon, Ajay; Murray, Christopher JL; Lauer, Jeremy A; Evans, David B*: Measuring Overall Health System Performance for 191 Countries. GPE Discussion Paper 30 (2001).
- [WHOR00] *WHO*: The World Health Report 2000 - Health System: Improving Performance. Geneva, Switzerland: World Health Organization, 2000.

Event analysis in Virtual Stock Markets for business research

Langfassung des Beitrags zum Doctoral Consortium der WI 2007

Dipl.-inform. Christian Slamka, M.Sc. (U.Conn.)

Professur für Electronic Commerce
Johann Wolfgang Goethe-Universität Frankfurt am Main
60054 Frankfurt
slamka@wiwi.uni-frankfurt.de

Hochschullehrer: Prof. Dr. Bernd Skiera

Professur für Electronic Commerce
Johann Wolfgang Goethe-Universität Frankfurt am Main
60054 Frankfurt
skiera@wiwi.uni-frankfurt.de

Abstract

Virtual Stock Markets (VSMs) are emerging as a valuable tool to forecast the outcome of future events, especially in the field of business research. While traditional methods in VSMs to perform business related predictions have concentrated on a static observation point, extended possibilities exist when taking into consideration the price dynamics and the price formation.

The dissertation aims at applying event analysis, i.e. the measurement of the events' effects resulting in stock price changes in VSMs, for the development of an application to forecast cannibalization effects of new and innovative products.

However, for being able to apply event analysis in VSMs, two further prerequisites have to be addressed.

First, in order for the events' effects to be incorporated in stock prices especially in small markets with low liquidity, an appropriate market maker will be developed. This trading mechanism provides participants with sufficient liquidity and enables them to react to events even in case of no matching counter offers.

The second prerequisite concerns the adoption of a methodology used in financial markets to evaluate events, the *event study*, to VSMs. In this case, mainly design variations in VSMs as well as the event study have to be considered.

1 Introduction and problem

Forecasting the outcome of future events has long been a field of research on its own. While many methods focus on past data and therewith trying to extrapolate future data, other methods, like expert interviews or surveys, do generally not rely on past data. A fairly novel method to predict future outcomes without reliance on past data are Virtual Stock Markets (VSMs). They have been proposed in the late 1980ies for political elections but are currently experiencing a constant upward trend in science as well as in application-oriented settings.

VSMs make outcomes to be forecasted tradable by modeling events' outcomes as stocks. E.g., in political elections, one stock could represent the shares of vote of the Democratic Party, the other one the shares of vote of the Republican Party. The stocks' final value (*payoff*) after the close of the market depends on the actual value of the outcome. In the current case, a the "Democratic" stock would be worth \$51 if the Democrats would have gained 51% shares of vote, where the "Republican" stock would be worth \$49 in case of 49% of votes. By trading the stocks on a market prior to the actual event, the participants' aggregated expectations about the events' outcomes are reflected in the stock price and consequently, serve as a predictions for the actual outcomes.

On the application side, VSMs have evolved from a tool to predict the outcomes of elections [FNNW92] to a wide variety of fields of application. These include, but certainly are not limited to sports forecasting (e.g. [ScWe02]) or political policy decision prediction (e.g. [WoZi06b]).

In business research, different but few applications for VSMs have been proposed so far. Considerably the most popular exchange in this area is the Hollywood Stock Exchange (HSX)¹, which offers unrestricted access to its markets and whose predictive accuracy has been determined to be very well [PLGN00; SpSk03]. On the other hand, in-house markets have been applied to forecast company-specific events. These include forecasting the completion time of a project as well as its duration [Ortn00], the forecasting of sales in a major IT company [KaCh02] or new product development [CDKL02; SoSp05]. Although without publishing their results in scientific articles, major companies such as Google² or Microsoft³ have started VSMs in in-house markets as well (also, c.f. [Suro04]).

Nearly all publications, with exception of more theory-oriented articles such as [PDGG02] or [DPGL03], do not take into account the dynamics, i.e. the movements, of prices in VSMs. Thus,

¹ <http://www.hsx.com>

² <http://googleblog.blogspot.com/2005/09/putting-crowd-wisdom-to-work.html>

³ c.f. <http://research.microsoft.com/~toddp/>

when making a prediction about a certain event, the only point in time which is being analyzed is shortly before the incidence of the event. Accordingly, a prediction about the outcome is derived at this particular moment.

However, studies and methods (so-called *event studies*) from financial markets (e.g., [CaLM97]) demonstrate that much more information is present in the dynamics of the market. Observing price developments of stocks and quantifying them has led to many new insights, as the high number of over 500 articles [KhWa06] in this area concerning financial markets shows. Examples of applications include the introduction of new products [ChDW91], the timing of stock option awards to CEOs [Yerm97] or the acquisitions of firms [KaWe92]. Thus, an event whose impact is usually hard to determine can be quantified with help of event studies by analyzing the abnormal returns the events cause.

From an application-oriented point of view, being able to measure effects of events in VSMs would be a great benefit, which can mainly be attributed to two reasons:

- In financial markets, events need to have a strong impact in order to be quantifiable. This implies that only large scale highly visible events can be measured. VSMs, on the other hand, can be specifically designed narrowly to forecast small scale outcomes and being able to react to small scale events.
- In VSMs, at least in laboratory environments, events can be created artificially. That is, during trading times, events such as news can be spread and reactions to them can be quantified.

At first glance, applying the event study approach to VSMs should not be a large problem, since VSMs apply the almost same market mechanism as financial markets. However, requirements for being able to conduct these studies have to be considered and met. Additionally, design variations of event studies have been tested in financial markets, but not yet in VSMs.

In very liquid markets participants are constantly willing to trade; however, this has not to be the case for small markets, where illiquidity can frequently be observed. This poses a problem because reaction upon events cannot occur with little liquidity in the market and thus, the events' effects cannot be reflected in the stocks' prices. In order to account for the problem, a source of liquidity is needed, providing participants with continuous sell and buy offers.

One actual problem which can be approached with event analysis is the forecasting of possible cannibalization effects of newly introduced products. In general, the main goal of introducing

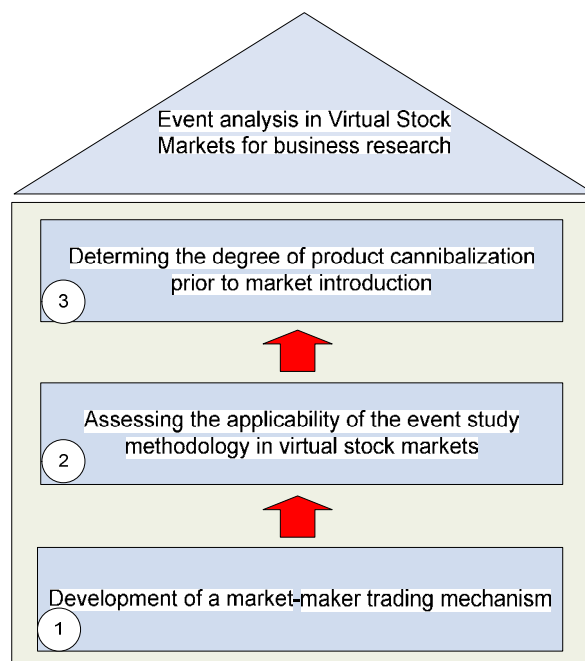
new, innovative products is to gain market share and therewith, raising revenues. If the new product takes market share from products already available on the market, so-called *cannibalization*, overall revenues of a brand suffer from this effect [Nijs97]. While this problem has been addressed with past data (e.g. [MaMi94]) by identifying cannibalization, this is hard to accomplish with innovative products because reliable past data is not present. In this respect, in a market simulation fashion, an approach of event analysis in VSMs is promising.

2 Goals of dissertation

The goal of this dissertation is to expand the potential use of VSMs for business forecasting by event analysis. In a first step (c.f. 4.1), an appropriate trading mechanism, an *automated market maker* (AMM) in this case, will be designed, implemented and tested. This AMM will serve as the base for efficient small scale markets.

In a second step (c.f. 4.2), the event study methodology will be transferred to VSMs and tested on two exchanges with different designs and structure.

With the preliminary work, an application is set up to analyze a real world business problem, for forecasting the cannibalization of products on a market (c.f. 4.3).



3 Virtual Stock Markets

By modeling future outcomes of forecasting goals as virtual stocks⁴, expectations of participants are made tradable. For each stock, an underlying payoff function determines the value of the stock after the close of the market. Thus, market participants speculate on the value of the outcome by using their available information to decide on trades. During the trading times, a rational trader will sell a stock if she assumes it to be overvalued, and will buy stocks, if she or he assumes it to be undervalued [WoZi06a]. Consequently, the price reflects the aggregate expectations of all participating traders in case of market efficiency [Fama70].

With unique information dispersed over possibly many participants (*private* information) and information present and available for all participants (*public* information), the price mechanism is the most effective instrument to aggregate asymmetrically dispersed information [Haye45].

In order to offer an incentive to reveal their true expectations, traders' remunerations are positively dependent on their performances, that is, their deposit values. Thus, the results of VSMs are less prone to be biased towards certain preferences or biases, compared to common polls or surveys (e.g. [SpSk03]).

The accuracy of the forecasts of PMs has been widely studied (e.g. [FNNW92; SpSk03]) and found to be generally at least as good as alternative forecasting methods such as opinion polls or expert surveys.

Various designs of markets have been used in the past. In general, the designs have to fit to the type of application and the environment the market is running in. Some of the most important design decisions, which will also be addressed either implicitly or explicitly in this dissertation, include the choice of

- the trading mechanism: usually a continuous double auction (CDA), but also automated market makers (AMMs) (c.f. 4.1),
- the payoff function: different variations exist, most important are linear, all-or-nothing and multi-part step functions (c.f. 3.1)
- the investment: a participant can either be endowed with play money or he can invest his own money.

⁴ Because of characteristics such as payoff functions, the term “option” is more suited, but because the term “stock” is used throughout the literature, it is adapted at this place.

From a pure IT perspective, designs for VSMs have already been discussed and implemented [Souk05]. Also, an implementation is already available as open source project⁵. Thus, from this point of view, does not seem to be an inherent need for further research at that moment, although advanced alternatives, e.g. providing trading functionality via web services, can always be considered.

3.1 Payoff functions

Not only is the choice of payoff functions important for prediction results [SpSk03], but also for further analyses, e.g., event studies (c.f. 4.2). These functions define the value of a stock after the close of the market, taking into account the outcome of the underlying event to be forecasted.

Three main types of payoff functions have prevailed and have been used in practice:

3.1.1 *Linear function*

In this case, the payoff is linearly dependent on the value of the outcome. Examples include the box-office revenues of movies, where the revenue is equal to the revenue divided by a certain constant (usually one million).

On the other hand, there are payoff functions whose values depend on the values of the outcomes on a relative basis, e.g. the shares of votes in an election where the percentage share of a candidate or party determines the value of a stock after the close of a market.

3.1.2 *All-or-nothing function*

In this case, the payoff is either a fixed value v (usually one hundred monetary units), or zero. Examples include the forecasting of champions in sports (“Team A’s stock will be worth \$100, if it wins the World Cup”).

The appeal in this function lies in the fact that stock prices reflect the probability of the occurrence of an event (which is the current price divided by the maximum payoff v).

3.1.3 *Multi-part step function*

Instead of utilizing a yes/no natured function, multi-part step functions often exist in case the number of stocks in a market is large.

⁵ C.f. <http://zocalo.sourceforge.net/>

Such kind of multi-part step functions are for instance frequently used in sports markets where teams are ranked in a table and for each position in the table, there is a corresponding payoff.

4 Approach

The dissertation consists of three main parts (depicted in Figure 1). Although the primary goal is to forecast cannibalization effects of products (part (3)), two main prerequisites have to be accomplished before dealing with the actual problem.

Firstly, the number of participants in VSMs can range from less than a dozen [Chri06] in small scale markets to hundreds of thousands (e.g., the HSX, with stocks issued for each new movie and actor). In company-internal or experimental markets, however, the problem of liquidity in markets plays a major role, which can be resolved by developing and applying a trading mechanism which is suited to act as mediator in low-liquidity markets (1).

The main methodology to be used to assess the impact of events will be the event study method (e.g., [CaLM97; McSi97]). In financial markets this method is commonly utilized, whereas in VSMs, it has not been systematically studied. Thus, the second part (2) tests for the applicability of this method in VSMs.

While in the second part, the focus is mainly on the event study methodology, in the last part, the method is applied to create a new alternative to assess cannibalization effects of new products.

4.1 Automated Market Maker

In financial markets, the method to match supply and demand is the trading mechanism. Markets can be categorized in two classes: *quote-driven* and *order-driven* (c.f. Figure 2).

In quote-driven markets or exchanges, such as the National Association of Securities Dealers Automated Quotations (*NASDAQ*), so-called *market makers (MMs)* state bid quotes and ask quotes, whereupon traders can accept the quoted prices. After a transaction (purchase or sale) has been completed, the price is adjusted, where the direction of adjustment is depending on the kind of transaction.

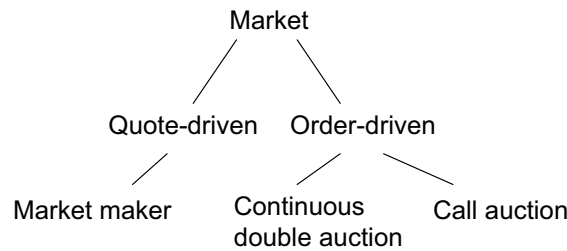


Figure 2: Classification of trading mechanisms

In order-driven markets and exchanges, such as the *New York Stock Exchange (NYSE)*, the price is solely determined by orders which are issued by market participants (sellers and buyers), thus, there is no intermediary such as a market maker.

With continuous double auctions (CDAs), buy (sell) limit orders are immediately executed if there is a matching sell (buy) limit order. Otherwise, the orders are stored in the order book for potential future execution. Thus, the CDA enables market participants in general to react to events and to adapt the stock prices accordingly. Also, so-called market orders exist; in this case, the order is immediately executed if a counteroffer exists.

In order to make trading in low-volume, order-driven markets possible, call auctions collect buy and sell offers for a fixed, pre-defined periods of time. At the time of execution, a price is fixed according to the principle of highest executable volume. Since during the time of order collection, prices are not determined, call auctions generally do not allow for continuous event evaluation.

While these order-driven trading mechanism in general work with VSMs and also have been implemented [Hans03; SpSk03], in small markets, there still persists the problem of illiquidity, intensified by “wrong” trading mechanisms. Illiquidity in order-driven markets occurs when transaction cannot take place in case of a lack of offers on either the buy or sell side. Especially in case of a large number of stocks on a market with few participants and asynchronous participation, illiquidity is a problem for the evaluation of impacts of events, which might not be possible due to missing counteroffers.

Therefore, few, exchanges in the field of VSMs like the HSX⁶ have implemented automated market makers (AMMs) for their exchange systems. However, no publication of their AMMs is available. In the field of market microstructure and empirical finance, some work without detailed explanations concerning AMMs exists [Madh92]. Recently, three papers were

⁶ Although the number of participants is high in the HSX, very illiquid stocks exist, making it necessary to provide a market maker.

published [Das05; Hans03; Penn04] dealing with AMMs in VSMs and financial markets. However, the suggestions do not meet the design requirements which are necessary in this work. In contrast to the AMMs mentioned in the three papers above, that compute different prices for bids and asks, this AMM only computes one price for both the bid and ask quotes. In this setting, there is no real money endowment and the AMM is not bound to make profits from the (usually existing) bid-ask spread.

In the model used in this part, which is comparable to the model of Glosten/Milgrom [GIMi85], there are informed traders as well as uninformed traders. Informed traders are exactly aware of the true value V of a stock. Thus, if the current price p is lower (higher) than V , an informed trader will buy (sell) the stock. Uninformed traders, on the other hand, do not possess information about the true value and will only act randomly. Hence, with informed traders mainly determining the price, the price in the equilibrium will converge to the true value V .

After a transaction has occurred, the price of the stock has to be adjusted by using a price adjustment mechanism, or price function, based on a set of past transactions [Das05; Hans03]:

$$f_i^{p_{\max}} : \mathbb{R}_0^+ \times T_i^I \rightarrow \mathbb{R}_0^+$$

where

i : index of current stock

p_{\max} : maximum price of stock

\mathbb{R}_0^+ : possible stock prices

T_i : transaction with $T_i = (p_i, q_i)$

I : size of window of past considered transactions

This function has to meet the following requirements:

- Purchases will lead to a price increase, sales to a price decrease.
- Higher quantity trades will affect the price more than lower quantity trades as they reflect higher certainty of the real value..
- The price is not higher than the maximum price p_{\max} , so that possible payoff restrictions can be met.

With different payoff functions in VSMs, the prices of stocks in a market are not independent. Thus, a price change of one stock will affect the other stocks since the sum of prices is fixed in an efficient market. Hence, in a second step, a price function for one stock will not only have to

affect the stock itself as seen in the formulations above, but also all other stocks in the market. Thus, the price function will change to

$$f_{i, \text{sum_payoffs}}^{P_{\max}} : (\mathbb{R}_0^+)^N \times T_i^I \rightarrow (\mathbb{R}_0^+)^N ,$$

where

N : number of stocks in the market,

sum_payoffs : the sum of payoffs.

Accordingly, the goal of this part of the dissertation is the development of the price function f . In order to test the price function, it is planned to conduct simulations as well as laboratory experiments. The simulations studies will be designed similar to those used in [Das05], with traders consisting of informed and uninformed traders in low liquidity market with continuously updated information flowing into the market. Besides the newly developed AMM, a CDA will be confronted to bring the results to a more valid basis. In the laboratory experiments, the AMM will be compared to a CDA.

4.2 Event Studies in VSMs

Since the beginning of the 1970s, event studies in financial markets have widely accepted as a standard procedure to measure the effect of an event on the firms' value, having proved usefulness in areas such as finance, accounting, law, economics or management.

By quantifying the abnormal change in the stock price, which specifies its value, before and after the announcements of the observed events, their effects can be measured. Subsequently, in management and application-oriented settings, implications for management and investors can be derived.

Although some methods and designs have been adopted from financial markets and have been applied in VSMs, this has not been the case for the event study methodology yet. As the evaluation of the impacts of events will be needed in the next step, the modified methodology for VSM is examined and the applicability of the method is tested.

Requirements for conducting an event study usually are not explicitly stated in application-oriented articles, but are implicitly assumed to be met. While this approach might be sufficient for financial markets, it has to be tested if the statistical and methodological requirements (stated by [McSi97]) can be met by VSMs:

- Market efficiency: All (publicly) available information has to be reflected in the stock (in general, security) price. In order to be able to conduct event studies, markets have to be at least semi-strong efficient.
- Confounding effects: It is crucial that no other events which might have an influence on the security price occur during the event study phase. Otherwise, the price reaction would not be completely attributed to the event.
- Unanticipated events: The events which occur are not incorporated into the security price prior to the event. Otherwise, measurement of the impact of the event would be inexact. If the event date is not clearly known, variations in the size of the event window can be made.

The procedure for event studies can roughly be described in words as follows. Firstly, the event and the event date to be analyzed have to be determined. Subsequently, two windows are defined, the *estimation window* and the *event window*. The data from the event window is used to determine the parameters of the model which is utilized to determine the performance of a stock, that is, the returns, if the event had not occurred. The estimation window starts before the actual event and precedes the event window, which captures the effects which are caused by the event. The excess, or *abnormal*, returns compared to the expected returns are aggregated in the event window. In order to make statistical statements about the validity of a possible change in returns, tests such as t-tests are applied to test the null hypothesis which states that returns of stocks have not changed as a cause of the impact of the event.

Throughout the literature, there exist a number of models to determine the expected returns of a certain stock. These are mainly classified into two categories: *statistical* and *economic* models [CaLM97]. However, in case of VSMs, two new classes are proposed: *market-based* models and *non-market-based* models. Market-based models take into account the performance of other stocks in the market, while non-market-based models do not. This is an important aspect when designing VSMs and choosing an appropriate payoff function. Depending on the payoff function, not every class of expected return model can be used. This is due to the fact that, depending on the payoff function, the prices of other stocks will be affected of by the price change of one stock (further details omitted). This leads to the following table:

		Payoff function				
		linear		all-or-nothing	multi-part step	
		relative values	absolute values	any	# single payoffs variable	# single payoffs fixed
Expected return model	non-market-based	✓	✓	✓	✓	✓
	market-based		✓		✓	

Table 1: Applicability of both return models and payoff functions

One important aspect when dealing with event studies in VSMs is the time period the markets are open for trading and the speed of the event sequence. In financial markets, event studies are usually conducted on data not finer grained than daily data. Therefore, an order-driven market has to be sufficiently liquid or a suited market maker has to be applied (c.f. 4.1).

In the empirical section of this part, the requirements stated above will be tested on two unique data sets (c.f. Table 2).

	Study 1	Study 2
Investment	real money	play money
# markets	31 (one per game)	1 (during championships)
# stocks	3 per game/market (Team A wins / Team B wins / draw)	16 (one for each team)
# participants	6.569	9.077
# transactions	88.174	1.108.293
Payoffs	Event occurs	100 € Champion
	Event does not occur	0 € Vice champion
		Elimination in semi-finals
		Elimination in quarter-finals
		3rd place group games
		4th place group games

Table 2: Data Sets

Another important test will be concerning market efficiency, using tests comparable to those in [GrGr97]. Further, tests concerning different event study modifications, such as various window sizes, will be conducted. When comparing the two exchanges, statements about the applicability of either payoff function or investment can be made and thus, consequences about market design when using event studies can be inferred.

4.3 Cannibalization

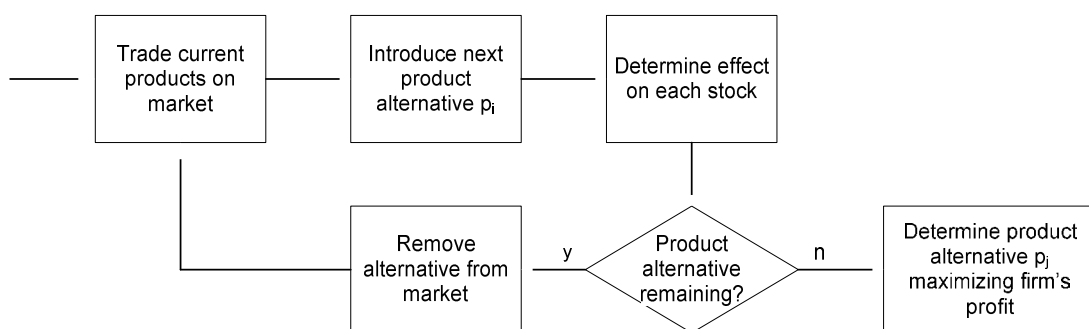
The key question of this part of the dissertation is to predict possible cannibalization effects before the introduction and development of new innovative products. Single, static predictions concerning new product success can, e.g., be made by modeling products as stocks [CDKL02; SoSp05]. The payoffs of each stock after the close of the market depend on the expected market share. These methods have proven external as well as internal validity and are currently further

being developed. In general, in this part of the dissertation, market dynamics are introduced and by quantifying the dynamics, i.e., price changes, statements about the impacts of newly created stocks can be made.

The basis idea (c.f. Figure 1) is to model an existing market (or market segment)⁷ by defining stocks on the market that correspond to the current products in the market. The payoff of each product is defined as the corresponding market share (a linear function, c.f. 3.1.1) and consequently, their payoffs, in case of an efficient market, prices, will add up to \$100.

After the equilibrium in the market has stabilized, a stock in form of an *initial public offering (IPO)* is introduced to the market. An IPO in financial markets is the initial listing of the stock on a market, where it is firstly traded in public. The introduction of product p_i will have an effect on the prices of the other stocks' prices, since the prices sum up to \$100. Thus, after the introduction, in case of market efficiency, the market prices will reflect the newly created situation including both the "old" stocks/products as well as the new stock.

The event (the introduction of the new stock)'s effects can be quantified by using the event study methodology whose applicability has been tested and described in part 4.2 of this proposal. After sufficient trading data for the analysis has been collected, the product alternative is taken off the market and trading can proceed with the old set of stocks. This procedure iterates until all product alternatives and their effects have been tested. After testing all alternatives, the alternative which maximizes the firm's profit is chosen among all other alternatives, especially considering the post-introduction phase, i.e., accounting for cannibalization effects which might occur to the same firm's products is considered.



While this is the general framework, questions for the design of the experiment and set-up remain:

- How can external validity be tested / proofed? One possible alternative is conjoint analysis.
- How can the problem of non-existing actual outcomes be attenuated [also addressed in CDKL02]?

5 Date of submission

The submission of the dissertation is planned for late 2008.

Literature

- [CaLM97] Campbell, John Y.; Lo, Andrew W.; MacKinlay, Craig M.: The Econometrics of Financial Markets. Princeton University Press, New Jersey 1997.
- [CDKL02] Chan, Nicholas T.; Dahan, Ely; Kim, Adlar; Lo, Andrew W.; Poggio, Tomaso: Securities Trading of Concepts (STOC). Technical report. Massachusetts Institute of Technology. (2002).
- [ChDW91] Chaney, Paul K.; Devinney, Timothy M.; Winer, Russel: The Impact of New Product Introductions on the Market Value of Firms. In: Journal of Business 64 (1991) 4, S. 573-610.
- [KaCh02] Chen, Kay-Yut; Plott, Charles R.: Information Aggregation Mechanisms: Concept, Design and Implementation for a Sales Forecasting Problem. Working paper. California Institute of Technology. (2002).
- [Chri06] Christiansen, Jed D.: Prediction markets; practical experiments in small markets and behaviours observed. In: Journal of Prediction Markets 1 (2006) 1, S. pp -.
- [Das05] Das, Sanmay: A Learning Market-Maker in the Glosten-Milgrom Model. In: Quantitative Finance 5 (2005) 2, S. 169-180.

- [DPGL03] Debnath, Sandip; Pennock, David M.; Giles, C. Lee; Lawrence, Steve: Information Incorporation in Online In-Game Sports Betting Markets. In: (Eds): Proceedings of the 4th ACM conference on Electronic commerce. ACM Press, San Diego, 2003, S. 258 - 259.
- [Fama70] Fama, Eugene F.: Efficient Capital Markets: A Review of Theory and Empirical Work. In: Journal of Finance 25 (1970), S. 383-417.
- [FNNW92] Forsythe, Robert; Nelson, Forrest; Neumann, George R.; Wright, Jack: Anatomy of an Experimental Political Stock Market. In: American Economic Review 82 (1992) 5, S. 1142-1161.
- [GIMi85] Glosten, Lawrence R.; Milgrom, Paul R.: Bid, Ask and Transaction Prices in a Specialist Market With Heterogeneously Informed Traders. In: Journal of Financial Economics 14 (1985), S. 71-100.
- [GrGr97] Gray, Philip K.; Gray, Stephen F.: Testing Market Efficiency: Evidence from the NFL Sports Betting Market. In: Journal of Finance 52 (1997) 4, S. 1725-1737.
- [Hans03] Hanson, Robin: Combinatorial Information Market Design. In: Information Systems Frontiers 5 (2003) 1, S. 107-119.
- [Haye45] Hayek, Friedrich August von: The Use of Knowledge in Society. In: American Economic Review 35 (1945) 4, S. 519-530.
- [KaWe92] Kaplan, Steven N.; Weisbach, Michael S.: The Success of Acquisitions: Evidence from Divestitures. In: Journal of Finance 47 (1992) 1, S. 107-138.
- [KhWa06] Khotari, S.P.; Warner, Jerold B.: Econometrics of Event Studies. In: Eckbo, B. E. (Eds): Handbook of Corporate Finance: Empirical Corporate Finance. Elsevier/North-Holland, 2006, S. Chapter 1.
- [Madh92] Madhavan, Ananth: Trading Mechanisms in Securities Markets. In: Journal of Finance 47 (1992) 2, S. 607-641.

- [MaMi94] Mason, C. H.; Milne, G. R.: An Approach for Identifying Cannibalization within Product Line Extensions and Multi Brand Strategies. In: Journal of Business Research 31 (1994), S. 163-170.
- [McSi97] McWilliams, Abigail; Siegel, Donald: Event Studies in Management Research: Theoretical and Empirical Issues. In: The Academy of Management Journal 40 (1997) 3, S. 626-657.
- [Nijs97] Nijssen, Edwin J.: Success factors of line extensions of fast-moving consumer goods. In: European Journal of Marketing 33 (1999) 5/6, S. 450-474.
- [Ortn00] Ortner, Gerhard: Aktienmärkte als Industrielles Vorhersagemodell. In: Zeitschrift für Betriebswirtschaft - Ergänzungsheft 70 (2000) 1, S. 115-125.
- [Penn04] Pennock, David M.: A Dynamic Pari-Mutuel Market for Hedging, Wagering, and Information Aggregation. ACM Conference on Electronic Commerce, New York, 2004.
- [PDGG02] Pennock, David M.; Debnath, Sandip; Glover, Eric J.; Giles, C. Lee: Modeling Information Incorporation in Markets, with Application to Detecting and Explaining Events. In: Darwiche, A.; Friedman, N. (Eds): Proceedings of the 18th Conference on Uncertainty in Artificial Intelligence (2002). Morgan Kaufmann Publishers, San Francisco, 2002, S. 405-413.
- [PLGN00] Pennock, David M.; Lawrence, Steve; Giles, Lee C.; Nielsen, Finn Arup: The Power of Play: Efficiency and Forecast Accuracy in Web Market Games. Technical report. NEC Research Institute. (2000).
- [ScWe02] Schmidt, Carsten; Werwatz, Axel: How Accurate Do Markets Predict the Outcome of an Event? The Euro 2000 Soccer Championship Experiment. Technical report. Max Planck Institut for Research into Economic Systems, Jena, Germany. (2002).
- [Souk05] Soukhoroukova, Arina: Flexible Software-Architektur für Prognosebörsen. Internationale Tagung Wirtschaftsinformatik, Bamberg, 2005, S. 231-250.

- [SoSp05] Soukhoroukova, Arina; Spann, Martin: New Product Development with Internet-based Information Markets: Theory and Empirical Application. 13th European Conference on Information Systems (ECIS), Regensburg, 2005.
- [SpSk03] Spann, Martin; Skiera, Bernd: Internet-Based Virtual Stock Markets for Business Forecasting. In: Management Science 49 (2003) 10, S. 1310-1326.
- [Suro04] Surowiecki, James: The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations. Doubleday, 2004.
- [WoZi06a] Wolfers, Justin; Zitzewitz, Eric: Interpreting Prediction Market Prices as Probabilities. Working paper. Stanford University GSB. (2006).
- [WoZi06b] Wolfers, Justin; Zitzewitz, Eric: Using Markets to Inform Policy: The Case of the Iraq War. Working paper. Stanford University GSB. (2006).
- [Yerm97] Yermack, David: Good Timing: CEO Stock Option Awards and Company News Announcements. In: Journal of Finance 52 (1997) 2, S. 449-476.

Kosten-Nutzen-Quantifizierung von Investitionen in IT-Sicherheit als Grundlage eines ökonomisch fundierten Sicherheitsmanagements

Thomas Nowey

Lehrstuhl Management der Informationssicherheit
Universität Regensburg
93040 Regensburg
thomas.nowey@wiwi.uni-regensburg.de

Abstract

Sicherheitsmanagement wird zunehmend als Risikomanagementaufgabe wahrgenommen. Verschiedene Faktoren erzwingen, dass Organisationen Sicherheitsrisiken bewerten und Kosten und Nutzen von Sicherheitsmaßnahmen abwägen. Hierzu sind quantitative Methoden unbedingt erforderlich. Während sich existierende Arbeiten häufig mit der Weiterentwicklung von Hilfsmittel zur Risikosteuerung und Entscheidungsunterstützung auf Basis quantitativer Methoden beschäftigen, entwickelt dieser Beitrag eine Möglichkeit solche quantitativen Daten überhaupt zu erheben. Hierzu wird ein System zur interorganisatorischen Sammlung von Informationen über Sicherheitsvorfälle konzipiert. Anforderungen an ein solches System werden dargestellt und Lösungsmöglichkeiten aufgezeigt. Abschließend werden künftige Herausforderungen, wie die Integration der gewonnenen Daten in ein umfassendes Sicherheitsmanagement, diskutiert.

1 Ausgangssituation und Problemstellung

Für das Sicherheitsmanagement wird zunehmend eine Orientierung an ökonomischen Prinzipien und damit verbunden eine Abwägung von Kosten und Nutzen der IT-Sicherheit – sowohl ex ante als auch ex post – gefordert (vgl. [SooH00], [Merc03], [Lubi06], [Pohl06]). Da auch für Sicherheitsinvestitionen die Annahme eines abnehmenden Grenznutzens sinnvoll

erscheint (vgl. [GoLo02a]), sind Unternehmen im Zuge von Effizienzüberlegungen auf der Suche nach dem optimalen Investitionsvolumen, aber auch nach Hilfsmitteln bei der Wahl der richtigen Alternativen. Externe Vorgaben (z. B. Basel II, Sarbanes-Oxley Act, KonTraG) verlangen, dass Unternehmen ihre Risiken laufend abschätzen und bewerten können und über geeignete Steuerungsinstrumente verfügen (vgl. [JoMü06]).

Seit 2002 wird die Problematik der Kosten-Nutzen-Evaluation häufig unter dem Schlagwort des Return on Security Investment (ROSI) diskutiert (vgl. [Beri02]), dessen Grundidee – analog zum klassischen ROI – der Vergleich des finanziellen Gewinns eines Investitionsprojekts mit seinen Gesamtkosten ist. Die Grundlage für den ROSI bildet das Konzept der jährlichen Verlusterwartung, Annual Loss Expectancy (ALE), welches bereits Ende der 1970er Jahre vom US National Bureau of Standards auf den Bereich der IT-Sicherheit übertragen wurde (vgl. [NBoS79]). Die erste Berechnungsformel für den ROSI wurde von Wissenschaftlern der University of Idaho im Rahmen einer Kosten-Nutzen-Analyse von Intrusion Detection Systemen angegeben:

$$\text{ROSI} = ALE_0 - ALE_1 - K \quad (1)$$

Wobei ALE_0 der jährlichen Verlusterwartung ohne eine Sicherheitsmaßnahme, ALE_1 der jährlichen Verlusterwartung mit dieser Sicherheitsmaßnahme und K den Kosten der Maßnahme entspricht. Inzwischen existiert eine große Zahl an Berechnungsvarianten (vgl. [Clus04]) und Erweiterungen des Basisansatzes. Verschiedenen Autoren haben bereits auf die Beschränkungen des Konzepts hingewiesen (vgl. z.B. [GoLo02b]). Zudem wurden verbesserte Verfahren für die Risikonalayse vorgeschlagen (vgl. [Scha06]).

Um derartige Kosten-Nutzen-Betrachtungen durchführen zu können, müssen Kosten und Nutzen von IT-Sicherheit messbar gemacht werden. Die Herausforderung liegt dabei vor allem im Bereich der Nutzenbestimmung. Aus dem ökonomischen Blickwinkel liegt der Nutzen von IT-Sicherheitsmaßnahmen hauptsächlich in der Reduktion von operativen Risiken. Aus Gründen der Vollständigkeit sei aber noch darauf hingewiesen, dass IT-Sicherheitsmaßnahmen auch anderweitig Erträge generieren können, z.B. durch die Vereinfachung betrieblicher Abläufe oder die Ermöglichung neuer geschäftlicher Aktivitäten. Sicherheitsmanagement ist also insbesondere eine Risikomanagementaufgabe (vgl. [Arba02], [RaDi04], [Stel02], [Loch05]). Diese Ausrichtung findet sich in diversen wissenschaftlichen Publikationen, ebenso wie in verschiedenen internationalen Standards wieder (neue ISO 27000er Reihe, NIST SP 800er Reihe, Méhari, BSI Grundschutzhandbuch).

Die meisten Veröffentlichungen konzentrieren sich dabei auf Ausschnitte des Risikomanagement-Kreislaufs (Identifikation, Bewertung, Steuerung, Überwachung). Ein anerkanntes Konzept zur ganzheitlichen Messung von Informationssicherheit existiert noch nicht ([MaST04]). Eine besondere Herausforderung stellt der Schritt der Risikobewertung dar. Um zu entscheiden, wie mit einem potentiellen Risiko umgegangen werden soll (vermeiden, reduzieren, abwälzen oder akzeptieren) ist es zunächst erforderlich, dieses möglichst genau abzuschätzen. Soo Hoo hat fünf zentrale Faktoren für ein quantitatives Risikomodell identifiziert ([SooH00]). Besonders schwer zu ermitteln, aber entscheidend für die Bewertung von Bedrohungen sind davon das potentielle Ausmaß der Bedrohung (Schadenshöhe) sowie andererseits die Häufigkeit des Ereignisses bzw. als analoge Größe dessen Eintrittswahrscheinlichkeit. Diese Werte werden bislang meist mit Hilfe qualitativer Daten (z.B. auf Basis von Checklisten oder Expertenurteilen) abgeschätzt und mit Hilfe ordinaler Skalen bewertet (vgl. z.B. OCATVE-Ansatz [AlDo02]). Für präzise Aussagen über die Risikolage eines Unternehmens sowie für ökonomisch sinnvolle Investitionsentscheidungen sind jedoch verlässliche, quantitative Daten nötig. Dies gilt auch für die häufig geforderten Versicherungen von Sicherheitsrisiken (vgl. [Schn02]).

2 Zielsetzung und Vorgehen

2.1 Zielsetzung

Gordon und Loeb konnten in ([GoLo06]) zeigen, dass inzwischen zahlreiche Unternehmen versuchen, ihre Investitionsentscheidungen im Bereich der IT-Sicherheit mit Hilfsmitteln aus der Investitionsrechnung zu stützen (z.B. Nettokapitalwert, interner Zinsfuß). In der Literatur wurden bereits auf den Bereich der IT-Sicherheit angepasste Weiterentwicklungen der oben genannten Methoden zur Risikosteuerung und Entscheidungsunterstützung vorgestellt (z.B. [FaPr06]). Um derartige Berechnungen durchführen zu können, müssen Ein- und Auszahlungsströme über mehrere Perioden hinweg möglichst genau prognostiziert werden. Ansätze zur quantitativen Analyse (vgl. [SooH00]) gehen bisher von der Annahme aus, dass entsprechende Daten bereits vorliegen und zeigen auf dieser Basis Wege zur Auswahl zwischen verschiedenen alternativen Investitionsprojekten auf. Befriedigende Lösungen für die

Ermittlung der nötigen Daten fehlen jedoch noch. Häufig bleibt zudem unklar, wie und auf welcher Ebene die Ergebnisse in das Sicherheitsmanagement einbezogen werden sollen.

Die Dissertation soll diese Lücke schließen. Ziel ist es, zu zeigen, wie solche Daten gewonnen und aufbereitet werden können und darzustellen, wie die so gewonnenen Informationen in ein umfassendes Vorgehen zum Sicherheitsmanagement integriert werden können. Zur Gewinnung quantitativer Daten kommen diverse Quellen in Frage. Die Sammlung historischer Daten hat sich in anderen Bereichen bereits bewährt, z.B. für Sach- und Lebensversicherungen. Einige Autoren argumentieren allerdings, historische Daten hätten für die Bewertung von Sicherheitsrisiken nur einen eingeschränkten Wert, da sich das Umfeld im Bereich der IT-Sicherheit rasant ändert und somit Vergangenheitsdaten nicht uneingeschränkt in die Zukunft extrapoliert werden können. Daher wurden marktbasierte Alternativen vorgeschlagen, die stärkeren Zukunftsbezug aufweisen sollen (vgl. [Ozme04], [Sche04], [Böhm06]). Die in diesen Artikeln diskutierten Instrumente existieren heute noch nicht und werden auch mittelfristig nicht zur Verfügung stehen. Auch wurde bislang auf theoretischer Ebene noch keine Abwägung vorgenommen, welche Datenquelle für welchen Einsatzzweck gewählt werden sollte. Die Arbeit soll eine solche Betrachtung durchführen. Darauf aufbauend ist ein System zum Austausch von Informationen über Sicherheitsvorfälle zwischen Organisationen zu entwickeln. Diverse Stellen sammeln bereits Informationen über Sicherheitsvorfälle (CERTs, Hersteller von Sicherheitssoftware). Das Ziel dabei ist aber meist die Früherkennung von Angriffen (z.B. Viren, Trojanische Pferde). Der hier propagierte Ansatz will solche Systeme weiterentwickeln und um Möglichkeiten zur Messung der Schadenshöhe und Eintrittswahrscheinlichkeit ergänzen. Der Autor möchte für die einzelnen oben dargestellten Problemfelder (Datengewinnung, Metriken, Vorgehensmodell) Lösungen entwickeln, diese evaluieren und in ein Gesamtkonzept integrieren. Ziel ist die Entwicklung einer Lösung, die mit bestehenden Standards zum Sicherheitsmanagement vereinbar ist und durch ihre Anpassbarkeit an verschiedene organisatorische Kontexte auch praktisch einsetzbar ist.

2.2 Vorgehen

Die Arbeit soll die Quantifizierung von Kosten und Nutzen der IT-Sicherheit umfassend strukturieren und diskutieren, um so zum Fortschritt auf dem Gebiet der ökonomischen Ausrichtung des Sicherheitsmanagements beizutragen. Ein besonderer Schwerpunkt wird dabei

auf die Ermittlung und Verwendung quantitativer Daten sowie auf das Zusammenführen technischer und ökonomischer Aspekte gelegt.

Der Arbeit liegt die Annahme zu Grunde, dass das Risikomanagement als wesentlicher Baustein der Kosten-Nutzen-Betrachtung für IT-Sicherheit gesehen werden kann (vgl.

ISO 27005). Die vier Schritte des Risikomanagementkreislaufs dienen als Grundlage für die einzelnen Untersuchungsschritte. Zunächst soll für jeden Bereich der Status Quo in Hinblick auf eine quantitative Kosten-Nutzen-Evaluation erhoben und analysiert werden. Anschließend werden die wesentlichen Handlungsfelder identifiziert und Teillösungen erarbeitet, die abschließend in ein Gesamtkonzept integriert werden.

Mögliche Quellen für quantitative Daten sollen identifiziert, verglichen und bewertet werden. Dabei soll insbesondere aufgezeigt werden, unter welchen Prämissen die Daten in der Praxis Verwendung finden können. Als Ausgangsbasis für weitere wissenschaftliche Untersuchungen soll ermittelt werden, welche quantitativen Daten Unternehmen benötigen, welche sie liefern können und wie sich diese ermitteln lassen.

Mit Hilfe von Literaturanalysen und empirischen Methoden (Experteninterviews, Fragebogen) soll ermittelt werden, wo genau der Informationsbedarf von Organisationen liegt und welche Daten zur Kosten-Nutzen-Evaluation bereits in heutigen Unternehmen vorliegen. Dabei sind geeignete Metriken und Kennzahlen zur Ermittlung der Parameter Schadenshöhe und Eintrittswahrscheinlichkeit zu entwickeln.

Durch die Entwicklung eines Prototyps kann evaluiert werden, inwieweit der Austausch von Informationen über Sicherheitsvorfälle zwischen Organisationen möglich ist und wie dadurch das Sicherheitsmanagement der einzelnen Organisation verbessert werden kann. Die Basis hierzu soll eine umfassende Analyse von Anforderungen an ein solches System darstellen. Dabei sollen insbesondere die Anforderungen in den Bereichen Sicherheit und Fairness analysiert und bewältigt werden.

Vorhandenen Vorgehensmodelle zum Sicherheitsmanagement werden auf ihre Erweiterbarkeit bezüglich einer quantitativen Kosten-Nutzen-Betrachtung untersucht und ggf. ergänzt. Das Vorgehensmodell soll sich dabei vor allem an Prozessen der Organisation orientieren und vorhandenes Wissen in der Organisation – z.B. aus dem Business Engineering – nutzen. Modelle aus der Entscheidungstheorie werden den verschiedenen Schritten des Risikomanagements zugeordnet und auf ihre Eignung in Bezug auf IT-Sicherheit überprüft.

Ross Anderson ([Ande01]) hat mit seinem Artikel "Why Information Security is Hard" gezeigt, dass ökonomische Theorien und Prinzipien auch im Bereich der IT-Sicherheit ihren Beitrag leisten können. Gordon, Loeb und Lucyshyn ([GoLL02]) haben so genannte ISACs (Information Sharing and Analysis Centres) aus mikroökonomischer Perspektive untersucht und einige Problemfelder und Erfolgsfaktoren identifiziert. Auch Gal-Or und Ghose ([GaGh04], [GaGh05]) haben Anforderungen aus mikroökonomischer Sicht an solche Systeme definiert und nachgewiesen, dass es zahlreiche Anreize für Firmen zur Teilnahme gibt. Diese Erkenntnisse aus der Mikroökonomie sollen zusammen mit Konzepten aus den Bereichen Investitionsrechnung und Controlling Eingang in die Arbeit finden.

3 Ein System zum überbetrieblichen Austausch von Informationen über Sicherheitsvorfälle

3.1 Grundkonzept

Historische Daten haben sich in anderen Bereichen, wie z.B. im Versicherungswesen, als Grundlagen für die Bewertung von Risiken bewährt. Nach Ansicht des Autors könnten sie auch für das Sicherheitsmanagement, entweder als direkte Schätzer oder als Grundlage für Simulationen, verwendet werden. Wie zu oben dargestellt, sind die relevanten Parameter, die für eine Risikobewertung benötigt werden, die Eintrittswahrscheinlichkeit und die mögliche Schadenshöhe. Diese können aus historischen Daten über Sicherheitsvorfälle abgeleitet werden. ISO/TR 18844 definiert einen Sicherheitsvorfall wie folgt: „An information security incident is indicated by a single or a series of unwanted or unexpected information security events that have a significant probability of compromising business operations and threatening information security.“. Zeichnet man Häufigkeit und Auswirkungen (Schaden) von Sicherheitsvorfällen auf, so lassen sich daraus die obigen Parameter ableiten, wobei die Eintrittswahrscheinlichkeit von Risiken aus dem wiederholten Eintreten eines bestimmten Vorfalls abgeleitet wird.

Nach Kenntnis des Autors gibt es bis heute keine allgemein anerkannte Datenbasis mit historischen Daten über IT-Sicherheitsrisiken. Natürlich zeichnen bereits verschiedene Organisationen Informationen über ihre Sicherheitsvorfälle auf und könnten diese Daten auch für den oben genannten Zweck verwenden, soweit sie auch entsprechende Informationen über

die entstandenen Schäden aufzeichnen. Gerade bei kleineren und mittleren Unternehmen bestünde hier jedoch die große Gefahr aufgrund einer zu geringen Datenbasis einzelne Risiken zu hoch oder zu niedrig einzuschätzen. Dies gilt insbesondere für so genannte HI/LF (High Impact/Low Frequency) Ereignisse, also solche Ereignisse, die zwar extrem selten auftreten, aber katastrophale Konsequenzen haben. Gerade diese Risikoart darf jedoch auf keinen Fall vernachlässigt werden, da sie den Fortbestand der gesamten Organisation gefährden kann. Daher ist nach Ansicht des Autors ein interorganisatorischer Ansatz nötig, der die Informationen aus verschiedenen Unternehmen aggregiert.

Insbesondere in den USA gab es schon in der Vergangenheit Initiativen die Unternehmen zum Austausch von Sicherheitsinvestitionen anregen sollten (vgl. [GoLL03]). Die Beteiligung an solchen Initiativen blieb aber vor allem aufgrund von Sorgen um den Ruf der Firma relativ gering. Eine explizit für diesen Zweck geschaffene, technische Plattform wurde dort aber nicht eingesetzt. Auch CERTs sammeln Daten von teilnehmenden Organisationen. Ihr Fokus liegt aber eher auf der Früherkennung und weniger auf der Sammlung von Informationen über Schadensereignisse.

Der Autor sieht daher die Notwendigkeit zur Schaffung eines Systems zum Austausch von Informationen über Sicherheitsvorfälle, welches seinen Schwerpunkt auf der Sammlung von Daten für das Risikomanagement hat. Gordon, Loeb und Lucyshyn haben sich in mehreren Papieren mit den mikroökonomischen Aspekten des Austausches von Sicherheitsinformationen beschäftigt (siehe [GoLL02],[GoLL03]). Die Autoren weisen darauf hin, dass Mechanismen, die das Fehlverhalten der Teilnehmer verhindern unbedingt erforderlich sind um eine solche Lösung praktisch einsetzbar zu machen. Insgesamt stellen sie fest, dass ein solcher Austausch einem Unternehmen helfen kann, sein optimales Sicherheitsniveau zu niedrigeren Kosten zu erreichen. Auch die gesamtwirtschaftlichen Wohlfahrtseffekte sind positiv. Dies unterstreicht den grundsätzlichen Nutzen einer solchen Plattform. Im Folgenden wird das Konzept für ein derartiges System genauer beschrieben.

Das System besteht aus den in Abb. 1 dargestellten Akteuren:

- Teilnehmer: Organisationen die innerhalb einer geschlossenen Benutzergruppe aktiv am Austausch von Vorfallsinformationen teilnehmen. Sie stellen Informationen über Sicherheitsvorfälle zur Verfügung und erhalten im Austausch dafür aggregiertes Datenmaterial für ihr Sicherheitsmanagement.

- **Platform Provider:** Eine Organisation die nicht aktiv am Informationsaustausch teilnimmt. Ihre Aufgabe besteht in der Bereitstellung der technischen Plattform, der Schaffung eines organisatorischen Rahmens und der Sammlung und Aufbereitung der Daten.
- **Externe Datenquellen:** Organisationen die zusätzliche Informationen bereitstellen. Diese können z.B. zur Erweiterung der Datenbasis oder zur Plausibilitätsprüfung übermittelter Vorfallsdaten verwendet werden.

Teilnehmende Organisationen sammeln und übermitteln Daten über jeden Sicherheitsvorfall der sich innerhalb ihrer Organisation ereignet. Innerhalb der Plattform werden die Daten aller Teilnehmer harmonisiert und gespeichert. Die Daten werden anschließend aggregiert, analysiert und interpretiert. Die Ergebnisse werden anschließend wieder an die Teilnehmer verteilt. Dies kann über eine Abfrageschnittstelle oder durch regelmäßige Reports erfolgen. Die Organisationen bekommen Informationen über Eintrittswahrscheinlichkeiten, Wahrscheinlichkeitsverteilungen und Erwartungswerte der Schäden und können so ihr Sicherheitsmanagement verbessern.

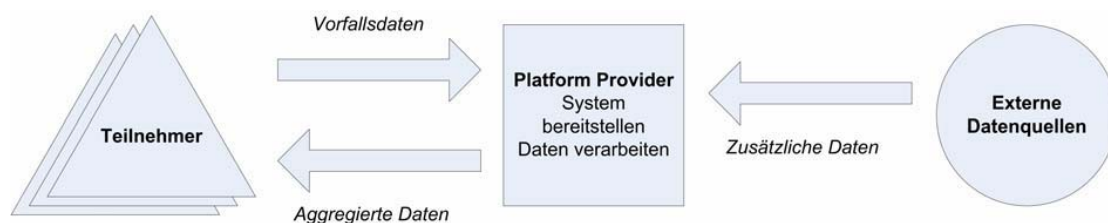


Abb. 1: Teilnehmer und ihre Aufgaben im Grundsystem

Abb. 2 illustriert die wesentlichen Aufgaben der zentralen Plattform:

- **Sammlung:** Das System stellt eine Schnittstelle zur Übertragung der Vorfallsdaten bereit und speichert die Datensätze. In diesem Schritt können auch Informationen aus externen Quellen berücksichtigt werden.
- **Harmonisierung:** Um eine Auswertung mit statistischen Methoden und eine Vergleichbarkeit der Daten zu gewährleisten, müssen die Vorfallsdaten standardisiert werden.
- **Aggregation:** Die standardisierten Daten werden anhand verschiedener Dimensionen zusammengefasst. So wird es beispielsweise möglich, das mehrfache

Auftreten eines Ereignistyps zu erkennen und Eintrittswahrscheinlichkeiten abzuleiten.

- Interpretation: In diesem Schritt werden die vorbereiteten Daten analysiert und ausgewertet. Werte für Eintrittswahrscheinlichkeiten, Erwartungswerte und deren Einflussfaktoren werden identifiziert und für die Nutzer bereitgestellt.

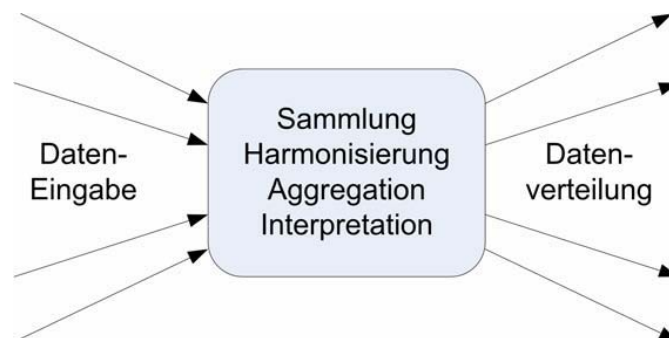


Abb. 2: Aufgaben der zentralen Plattform

Als wichtigste Zielgruppe für das beschriebene System werden kleine und mittlere Unternehmen (KMU) angesehen, da diese sehr eingeschränkte Möglichkeiten haben selbst Daten zu sammeln. Die einzige Alternative für diese Organisationen wäre regelmäßig teure Expertenmeinungen einzukaufen. Erste Experteninterviews zeigten, dass bei KMU die Sicherheit nicht als zentralen Wettbewerbsvorteil sehen, Bereitschaft besteht, Daten für ein solches System zur Verfügung zu stellen, um damit die Effizienz ihres Sicherheitsmanagements zu verbessern, wenn bestimmte Bedingungen erfüllt sind (vgl. nächster Abschnitt).

Bereits bei der Darstellung des Basissystems wurde daraufhin gewiesen, dass das Funktionieren eines solchen Systems an die Erfüllung gewisser Anforderungen gebunden ist. Diese werden im nächsten Abschnitt vorgestellt.

3.2 Anforderungen

3.2.1 Sicherheit

Die Informationen die innerhalb des Systems ausgetauscht werden sind sehr sensibel. Daher ist Sicherheit eine zentrale Anforderung an die Plattform. Gemäß den Prinzipien der mehrseitigen Sicherheit, müssen die Schutzinteressen aller beteiligten Parteien berücksichtigt werden ([FePf97]). Die an das System übertragenen Vorfallsdaten verraten zahlreiche Informationen

über die Teilnehmer, wie z.B. Sicherheitslücken, Sicherheitsmaßnahmen, usw. Eine Enthüllung dieser Informationen könnte Angreifern wertvolle Hinweise geben. Außerdem fürchten Unternehmen negative Wirkungen auf ihren Ruf, wenn Informationen über Sicherheitsvorfälle bekannt werden. Daher ist das Schutzziel Vertraulichkeit besonders wichtig. Da die Informationen, die von der Plattform geliefert werden, direkten Eingang in das Sicherheitsmanagement der Organisationen finden, ist auch die Integrität der Daten von großer Wichtigkeit. Verfälschte Daten könnten zu suboptimalen Entscheidungen führen. Weiterhin basiert das System auf einer geschlossenen Nutzergruppe. Der Plattform Provider muss durch geeignete Mechanismen sicherstellen, dass nur berechtigte Nutzer Daten einstellen und abrufen können.

Im Folgenden werden die wesentlichen Ziele im Bereich Sicherheit, aufgeteilt nach den drei Schutzzielen dargestellt:

Vertraulichkeit:

- Vertraulichkeit der Kommunikation zwischen Plattform und Teilnehmern.
- Vertraulichkeit der Daten in der Datenbasis. Lediglich dem Plattform Provider muss ein Zugriff möglich sein.
- Unverkettbarkeit zwischen Teilnehmern und Sicherheitsvorfällen.
- Unbeobachtbarkeit der Kommunikation zwischen Teilnehmern und Plattform.

Integrität:

- Integrität der Kommunikation zwischen Plattform und Teilnehmern.
- Integrität der Daten innerhalb der Plattform
- Authentifizierung und Autorisierung der Teilnehmer.

Verfügbarkeit:

Da es sich nicht um ein Echtzeitsystem handelt, ist Verfügbarkeit in diesem Fall nicht so wichtig wie die anderen beiden Schutzziele. Dennoch sollte sichergestellt werden, dass das System für die Teilnehmer nutzbar ist.

Die Nutzung folgender Sicherheitsmechanismen ermöglicht die Erreichung der Schutzziele:

- Einsatz einer Trusted Third Party. Dies könnte aufgrund seiner ohnehin zentralen Rolle vom Platform Provider übernommen werden. Es sollte sich folglich um eine unabhängige Institution handeln, die nicht selbst am Informationsaustausch teilnimmt.
- Einsatz kryptographischer Mechanismen zur Sicherstellung von Vertraulichkeit und Integrität (inklusive Authentifizierung).
- Einsatz datenschutzfreundlicher Techniken, insbesondere von Mechanismen zur Anonymisierung der Kommunikation zwischen Teilnehmer und Plattform sowie der Einsatz von Pseudonymen für die Teilnehmer. Dabei ergeben sich möglicherweise Konflikte mit dem Ziel Fairness.

3.2.2 *Fairness*

Eng verknüpft mit dem vorherigen Abschnitt ist der Aspekt der Fairness. Damit sind in diesem Kontext vor allem die folgenden drei Probleme gemeint:

- Free-riding: Einzelne Organisationen könnten sich als Teilnehmer an der Austauschplattform beteiligen, ohne jemals selbst Informationen zur Verfügung zu stellen. Solches Verhalten kann den Erfolg der gesamten Plattform gefährden. (siehe [GoLL03])
- Übermittlung falscher Daten: Organisationen könnten bewusst falsche Daten übermitteln, um die anderen Teilnehmer zu schädigen.
- Datenmissbrauch: Organisationen könnten Daten aus dem System an Dritte weitergeben oder selbst für Angriffe nutzen.

Vier Typen von Maßnahmen könnten zur Gewährleistung von Fairness beitragen:

- Anreizsysteme: Ein solches System könnte beispielsweise das Übermitteln von Vorfallsinformationen belohnen.
- Reputationssysteme: Ausgewählte Datensätze der Teilnehmer werden anderen Teilnehmern zur Bewertung vorgelegt.
- Vertragliche Regelungen: Solche Regelungen können zwar Missbrauch nicht verhindern, aber zumindest klare Regeln aufstellen und beispielsweise Konventionalstrafen androhen.

- Technische Mechanismen: Denkbar sind z.B. automatische Plausibilitätskontrollen oder der Vergleich der übermittelten Informationen mit Daten aus externen Quellen.

3.2.3 *Gemeinsame Sprache und Metriken*

Sowohl die Verwertbarkeit innerhalb der Plattform als auch die Nutzbarkeit der Daten durch die Teilnehmer werden entscheidend von der Qualität der Daten beeinflusst. Einerseits kann das Ziel des Plattform Providers, Daten zu aggregieren und auszuwerten, nur erreicht werden, wenn die unterschiedlichen Vorfallsbeschreibungen zueinander kompatibel sind. Andererseits muss es einer Organisation möglich sein die erhaltenen Informationen auf ihren eigenen Kontext zu übertragen. Um die Vergleichbarkeit und Übertragbarkeit der Daten zu begünstigen, bieten sich folgende Mechanismen an:

- Eine gemeinsame Sprache für Sicherheitsvorfälle. Der gleiche Sicherheitsvorfall muss von unterschiedlichen Organisationen gleich beschrieben werden. Hier bietet sich die Verwendung einer Taxonomie an, die eine eindeutige Zuordnung ermöglicht (z.B. [HaHu05]).
- Einheitliche Metriken für die Beschreibung von Schäden. Da alle Organisationen unterschiedlich sind, wäre es nicht präzise, Schadensdaten verschiedener Organisationen einfach zusammenzufassen. Stattdessen werden gemeinsame Nenner für die einzelnen Dimensionen zur Charakterisierung von Schäden benötigt.

3.2.4 *Vorgehensmodell*

Wie eingangs erläutert, existiert nach Wissen des Autors kein umfassendes Vorgehensmodell zum Sicherheitsmanagement, welches Aufschluss darüber gibt, wie quantitative Daten in das Sicherheitsmanagement integriert werden können. Ein solches Vorgehensmodell, sollte den Nutzer bei der Auswahl geeigneter Metriken und Risikosteuerungsinstrumente unterstützen und aufzeigen, wie Standards und regulatorische Auflagen mit Hilfe der quantitativen Daten erfüllt werden können.

3.2.5 *Usability*

Die Qualität der Ausgabedaten steigt mit der Anzahl der übertragenen Vorfallsdaten. Daher muss es den Teilnehmern so einfach wie möglich gemacht werden, diese Informationen an das

System zu übertragen. Dies könnte durch ein leicht zu bedienendes Web-Interface für die manuelle Eingabe erreicht werden. Andererseits sollten auch Schnittstellen für Organisationen bereitgestellt werden, die bereits andere Applikationen zur Behandlung von Vorfallsdaten einsetzen.

3.2.6 Interaktionsmöglichkeiten

Eine Anforderung, die sich aus der Diskussion mit potenziellen Nutzern ergeben hat, ist der Wunsch nach Möglichkeiten zur Interaktion zwischen den Teilnehmern. Neben genauen statistischen Daten, sollte auch die Möglichkeit zum Austausch über Erfahrungen und Best Practice bestehen. Dies erfordert natürlich einerseits den Aufbau von Kommunikationsbeziehungen zwischen den Teilnehmern und andererseits die Veröffentlichung von Informationen in nicht aggregierter Form. Hier entsteht ein Zielkonflikt mit dem Wunsch nach Schutz der Nichtzuordenbarkeit von Vorfällen zu Unternehmen. Deshalb sollte es jedem Teilnehmer selbst überlassen werden, zu entscheiden welche Informationen er auch anderen Organisationen zugänglich machen möchte.

4 Fazit und weiteres Vorgehen

Im vorliegenden Papier wurde ein Weg zur Verbesserung der Kosten-Nutzen-Quantifizierung von Investitionen in IT-Sicherheit aufgezeigt. Verschiedene zu bearbeitende Aspekte wurden identifiziert und entsprechende Methoden dargestellt.

Die zentrale Idee einer Plattform zur Sammlung von Daten über Sicherheitsvorfälle wurde erläutert. Auch wurde gezeigt, dass eine Plattform grundsätzlich für die Teilnehmer vorteilhaft ist. Dies geschieht allerdings nur unter der Voraussetzung, dass verschiedene Rahmenbedingungen eingehalten werden. Diese wurden dargestellt und Lösungswege aufgezeigt. Besonders wichtig erscheint die Gewährleistung von Sicherheit und Fairness. Während sich die Gewährleistung der verschiedenen Schutzziele im Bereich Sicherheit gut umsetzen lässt, ist im Bereich der Anreizsysteme für faires Verhalten der Systemteilnehmer noch Forschungsbedarf gegeben.

Ziel der zu entwickelnden prototypischen Plattform zum Austausch von Informationen über Sicherheitsvorfälle ist es, einen Weg aufzeigen, wie quantitative Daten zu relevanten Schadensparametern aggregiert und vergleichbar gemacht werden können. Die so gewonnenen

Daten können in vielfältiger Weise genutzt werden. So z.B. im wissenschaftlichen Bereich zur Evaluation bestehender Sicherheitsmanagementansätze oder zur Verfeinerung von ökonomischen Modellen im Bereich der IT-Sicherheit. Organisationen in der Praxis können mit Hilfe dieser Daten ökonomisch fundierte Entscheidungen treffen. Hierzu können die Daten direkt in Ansätze zur Risikobewertung und –steuerung übernommen werden oder aber als Input für Simulationen dienen. Schließlich können solche Daten auch einen Beitrag für Wissenschaft und Praxis im Bereich der Versicherung und Versicherbarkeit von IT-Sicherheitsrisiken leisten. Mit einer ausreichend großen Datenbasis könnten einerseits existierende Produkte optimiert werden, andererseits könnte ein Beitrag zur Entwicklung neuer Produkte geleistet werden, da so z.B. Hypothesen bezüglich Ereigniskorrelationen oder Verteilungsfunktionen überprüft werden könnten.

Ein weiteres wichtiges Ziel ist es aufzuzeigen, wie die über die oben dargestellte Plattform gewonnen Informationen in das Sicherheitsmanagement einer Organisation integriert werden können. Dabei soll insbesondere auf eine Kombinierbarkeit mit fortgeschrittenen Ansätzen der Risikosteuerung geachtet werden. Durch die Integration der ermittelten Teilresultate in ein Gesamtverfahren erhalten Organisationen ein praktisch anwendbares Vorgehensmodell für die Kosten-Nutzen-Evaluation von Sicherheitsinvestitionen. Die Quantifizierung von Kosten und Nutzen ermöglicht nicht nur verbesserte Investitionsentscheidungen sondern hilft Unternehmen auch externe Anforderungen zu erfüllen.

Die nächsten Schritte bestehen nun in der Entwicklung von Mechanismen zur Gewährleistung von Fairness innerhalb der Plattform und in der Identifikation potenzieller Systemnutzer, welche für empirische Untersuchungen und den Test des prototypischen Systems zur Verfügung stehen.

Literaturverzeichnis

- [AlDo02] Alberts, Christopher J.; Dorofee, Audrey J.: Managing Information Security Risks: The OCTAVE™ Approach. Addison-Wesley, New York 2002.
- [Ande01] Anderson, Ross J.: Why Information Security is Hard - An Economic Perspective. In: ACSAC, 2001, S 358–365.

- [Arba02] Arbaugh, William A.: Security: Technical, Social, and Legal Challenges. In: IEEE Computer, 35(2) 2002, S. 109–111.
- [Ber02] Berinato, Scott: Finally a Real Return on Security Spending. In: CIO Magazine, 2002.
- [Böhm06] Böhme, Rainer: A Comparison of Market Approaches to Software Vulnerability Disclosure. In: Emerging Trends in Information and Communications Security. International Conference, ETRICS 2006, Freiburg, Germany, June 2006. Proceedings 2006, S. 298-311.
- [Clus04] Clusif Groupe de Travail ROSI (Hrsg.): Retour sur investissement en s'écrit'e des systèmes d'information: Quelques clés pour argumenter, Oktober 2004.
- [FePf97] Federrath, Hannes; Pfitzmann, Andreas: Bausteine zur Realisierung mehrseitiger Sicherheit. In: Günter Müller, Andreas Pfitzmann (Hrsg.): Mehrseitige Sicherheit in der Kommunikationstechnik, Addison-Wesley-Longman 1997, S. 83–104.
- [FaPr06] Faisst, Ulrich; Prokein, Oliver: Management of security risks - a controlling model for banking companies. http://www.wiwi.uni-augsburg.de/bwl/buhl/dyn/root_wissenschaft/030Publikationen/pdf/wi-170.pdf, Abruf am 30.11.2006.
- [GaGh04] Gal-Or, Esther; Ghose, Anindya: The Economic Consequences of Sharing Security Information. In: The Economics of Information Security, Kluwer 2004, S. 95–104.
- [GaGh05] Gal-Or, Esther; Ghose, Anindya: The Economic Incentives for Sharing Security Information. In: Information Systems Research, 16(2) Juni 2005, S. 186–208.
- [GoLo02a] Gordon, Lawrence A.; Loeb, Martin P.: The Economics of Information Security Investment. In: ACM Transactions on Information and System Security, 5(4), November 2002, S. 438–457.

- [GoLo02b] Gordon, Lawrence A.; Loeb, Martin P.: Return on Information Security Investments: Myths vs. Reality. In: Strategic Finance, 84(5), November 2002, S. 26–31.
- [GoLo06] Gordon, Lawrence A.; Loeb, Martin P.: Budgeting Process for Information Security Expenditures. In: Communications of the ACM, 49(1), Januar 2006, S. 121–125.
- [GoLL02] Gordon, Lawrence A.; Loeb, Martin P.; Lucyshyn, William: An Economics Perspective on the Sharing of Information Related to Security Breaches: Concepts and Empirical Evidence. In: Proceedings of the First WEIS 2002.
- [GoLL03] Gordon, Lawrence A.; Loeb, Martin P.; Lucyshyn, William: Sharing information on computer systems security: An economic analysis. In: Journal of Accounting and Public Policy, 22(6) 2003, S. 461–485.
- [HaHu05] Hansman, Simon; Hunt, Ray: A taxonomy of network and computer attacks. In: Computers & Security 24(1) 2005, S. 31–43.
- [JoMü06] Jonen, Andreas; Müller, Jochen: Risk-Evaluator - Risikoorientierte Entscheidungsunterstützung bei Investitionen in IT. In: Lehner, Franz; Nösekabel, Holger; Kleinschmidt, Peter (Hrsg.): Multikonferenz Wirtschaftsinformatik 2006, Band 2, S. 141–154.
- [Loch05] Locher, Christian: Methodologies for evaluating information security investments – What Basel II can change in the financial industry. In: Proceedings of the 13th European Conference on Information Systems, Information Systems in a Rapidly Changing Economy, ECIS 2005, Regensburg, Germany, May 26–28, 2005 .
- [Lubi06] Lubich, Hannes P.: IT-Sicherheit: Systematik, aktuelle Probleme und Kosten-Nutzen-Betrachtungen. In: HMD, Praxis der Wirtschaftsinformatik, (248), April 2006, S. 6–15.

- [MaST04] Matousek, Michael; Schlienger, Thomas; Teufel, Stephanie: Metriken und Konzepte zur Messung der Informationssicherheit. In: HMD, Praxis der Wirtschaftsinformatik, (236) 2004, S.33-41.
- [Merc03] Mercuri, Rebecca T.: Analyzing Security Costs. In: Communications of the ACM, 46(6) 2003, S. 15-18.
- [NBoS79] National Bureau of Standards (Hrsg.): Guideline for Automatic Data Processing Risk Analysis, FIPS PUB 65 1979.
- [Ozme04] Ozment, Andy: Bug Auctions: Vulnerability Markets Reconsidered. In: Workshop of Economics and Information Security (WEIS). Minneapolis, MN, 2004. <http://www.dtc.umn.edu/weis2004/ozment.pdf>, Abruf am 30.11.2006.
- [Pohl06] Pohlmann, Norbert: Wie wirtschaftlich sind IT-Sicherheitsmaßnahmen. In: HMD, Praxis der Wirtschaftsinformatik, (248), April 2006, S. 26–34.
- [RaDi04] Rauschen, Thomas; Disterer, Georg: Identifikation und Analyse von Risiken im IT-Bereich. In: HMD - Praxis der Wirtschaftsinformatik, Nr. 236, 2004, S. 19-32.
- [Scha06] Schadt, Dirk: Über die Ökonomie der IT-Sicherheit. In: HMD, Praxis der Wirtschaftsinformatik, (248), April 2006, S. 16–25.
- [Sche04] Schechter, Stuart E.: Computer Security Strength & Risk: A Quantitative Approach. Dissertation, Harvard University, Cambridge, MA 2004.
- [Schn02] Schneier, Bruce: It's the economics stupid, 2002. <http://www2.sims.berkeley.edu/resources/affiliates/workshops/econsecurity/econws/18.doc>, Abruf am 30.11.2006.
- [SooH00] Soo Hoo, Kevin J.: How Much Is Enough? A Risk-Management Approach to Computer Security, 2000. <http://iis-db.stanford.edu/pubs/11900/soohoo.pdf>, Abruf am 30.11.2006.

- [Stel02] Stelzer, Dirk: Risikoanalysen als Hilfsmittel zur Entwicklung von Sicherheitskonzepten in der Informationsverarbeitung. In: IT-Sicherheitsmanagement in Banken, Bankakademie-Verlag 2002, S. 37–54.

Analyse der Selbstorganisation in Wikis

Claudia Müller

Lehrstuhl Wirtschaftsinformatik und Electronic Government
Universität Potsdam
14882 Potsdam
{cmueller}@wi.uni-potsdam.de

Abstract

Im betrieblichen Wissensmanagement finden Wikis zunehmend Anwendung, um virtuelle Wissensgemeinschaften informationstechnisch zu unterstützen. Wikis wird nachgesagt, die Prozesse der Selbstorganisation zu fördern. In der wissenschaftlichen Diskussion bleibt aber offen, was Selbstorganisation in diesem Zusammenhang bedeutet und wie sie nachgewiesen werden kann. In dem folgenden Beitrag wird ein Ansatz beschrieben, die Netzwerkstruktur in Wikis auf Effekte der Selbstorganisation zu untersuchen.

1 Einführung

Dieser Beitrag stellt den aktuellen Stand eines an der Universität Potsdam durchgeführten Promotionsvorhabens vor. Zunächst wird die Motivation zur Durchführung dieser Promotion verdeutlicht. Es erfolgt die Beschreibung des Untersuchungsgebiets, mit dessen wesentlichen Eigenschaften, insbesondere der Selbstorganisation. Aus dieser Eigenschaft leitet sich der zweite Teil des Beitrags ab, in welchem die Forschungsfrage erläutert wird. Im dritten Teil wird ein Beschreibungsansatz der Selbstorganisation, die Synergetik, auf das Untersuchungsgebiet übertragen, um das gewählte Vorgehen zu verdeutlichen. Im letzten Teil des Beitrags werden die nächsten Arbeitsschritte kurz skizziert.

1.1 Netzwerke in Unternehmen

Als John Naisbitt 1982 in seiner wirtschaftlichen, politischen sowie gesellschaftlichen Analyse der US-amerikanischen Gesellschaft zukünftige Megatrends postulierte, konnte er noch nicht ahnen, dass diese Trends auch für andere Länder gelten [Nai82]. Heute, 25 Jahre

später, haben sich viele seiner Vorhersagen bewahrheitet. Die Verflechtung und gegenseitige Abhängigkeit der Weltwirtschaft hat im Zuge des Wandels von national zu global handelnden Unternehmen zugenommen. Es hat ein Übergang von der Industrie- zur Informationsgesellschaft stattgefunden, in welcher hoch entwickelte Technologien einen bedeutenden Platz einnehmen. Zentrale, hierarchische Organisationsformen entwickelten sich zu dezentralen und vernetzten Strukturen, mit verstärkter vertikaler und horizontaler Kooperation. In diesen offenen, vernetzten Systemen sind die zentralen Ressourcen der Mensch und die verfügbaren Informationen [TaCa93].

Offene Systeme sind einem kontinuierlichen Veränderungsprozess unterworfen. Während es für hierarchische Organisationen aufgrund der fehlenden Entscheidungsflexibilität schwierig ist, Anforderungsänderungen in die laufenden Prozesse zu integrieren, fällt das Netzwerkorganisationen um ein Vielfaches leichter, da das Gestaltungs- und Lenkungspotential redundant auf das gesamte Unternehmen verteilt ist (vgl. [Bar03]). Fest definierte Arbeitsprozesse nehmen ab, Mitarbeiter werden zunehmend flexibel aufgabenbezogen eingesetzt und ihnen wird mehr Eigenverantwortung und Selbststeuerung abverlangt [Sco98]. Unternehmen verfügen zwar über ein eindeutiges Zentrum, aber ihre Prozesse sind verstärkt dezentralisiert. Flache Hierarchien und flexible Strukturen, Teamorientierung sowie die steigende Abhängigkeit vom Wissen sind wesentliche Beschreibungsmerkmale heutiger Unternehmen [Cros02].

1.2 Bildung von virtuellen Wissensgemeinschaften

Diese Veränderungen haben Auswirkungen auf das betriebliche Wissensmanagement. Zur Sicherstellung des organisationalen Lernens sollte die Bildung von Wissensgemeinschaften gefördert werden [Rom02]. Diese, nicht-strukturell legitimierten Wissensverknüpfungen einer Organisation, wurden zunächst im Rahmen der Communities of Practices (CoP) Forschung analysiert (vgl. [Wen98]). Die CoP-Forschung basiert auf sozialen Lerntheorien, in welchen das Lernen ein selbstorganisierter und selbststeuernder Prozess von Personengruppen ist. Es kann zwar ein für das Lernen förderndes Umfeld geschaffen werden, aber dieses muss nicht von außen organisiert sein [Der99]. Wissen ist personengebunden und dynamisch, es verändert sich stetig, abhängig von den Erfahrungen einer Person. Es ist nicht abgrenzbar und isoliert assimilierbar [Neu99]. Vielmehr entsteht Wissen in einem Prozess der Interaktion von Personen in unterschiedlichen Kontexten [NoRP00].

Für das betriebliche Wissensmanagement sollten Rahmenbedingungen geschaffen werden, damit Wissensgemeinschaften entstehen, in welchen Wissen generiert und ausgetauscht wird. Das sollte mit dem Ziel verbunden sein, dieses Wissen in die Wertschöpfungsprozesse des Unternehmens zu integrieren. Auch wenn dies nicht originär im Konzept der Wissensgemeinschaften von Romhardt verankert ist [Rom02], so benötigen verteilt agierende Mitarbeiter (siehe Abschnitt 1.1) informationstechnische Hilfsmittel. Eine Möglichkeit dieser Unterstützung bieten Wiki-Anwendungen [Klo06]. Im Gegensatz zu „herkömmlicher“ Software stehen bei Wikis nicht die individuelle Produktivitätsorientierung im Vordergrund, sondern die Gestaltung von Beziehungen zwischen Personen oder Personengruppen in Netzwerken [HiWi05]. Das erste Wiki wurde 1995 von Ward Cunningham als Portland Pattern Repository implementiert [LeCu01]. Allgemeine Bekanntheit erlangte diese Software aufgrund des Wikipedia-Projekts.

Aufgrund der thematischen Nähe ist ein Haupteinsatzgebiet für Unternehmens-Wikis die Softwareentwicklung. Hier werden Wikis zur kollaborativen Erstellung von Dokumentationen, zur Erfassung von Softwarefehlern (nur intern oder als Kommunikationsinstrument mit dem Kunden), aber auch zur Ideensammlung, beispielsweise zur Erfassung von Vorgehensweisen bei der Programmierung eingesetzt ([BiMu06], [MaWY06], [Klo06]). Ebenfalls liefern Wikis einen wertvollen Beitrag zur Koordination der Projektaktivitäten. Wichtige Projektinformationen werden für alle verfügbar und direkt erweiterbar in das Wiki gestellt. Beispielsweise entfällt das Verschicken von Besprechungsprotokollen per E-Mail, da das Protokoll im Wiki abgelegt wird. Im Nachhinein erfolgte Änderungen können nachvollzogen werden. Das Zusammentragen von Besprechungsthemen kann aufgrund der zentralen Verfügbarkeit über das gesamte Team erfolgen. Auch für das persönliche Wissensmanagement sind Wikis geeignet.

1.3 Funktionen eines Wikis für das Wissensmanagement

Der Informationsraum eines Wikis besteht aus Artikeln, die assoziativ über Relationen verknüpft sind. Diese Verlinkung erzeugt eine netzartige Struktur, aufgebaut durch die Verbindungen (Hyperlinks) zwischen den Inhalten, d.h. Informationseinheiten (Text, Textteile) und/oder Objekten (Dateien). Die mathematische Beschreibung von Netzwerken erfolgt mit Hilfe eines Graphs, der aus einer endlichen Menge von Knoten und Kanten besteht. Da Wissen dynamisch ist und deshalb häufig aktualisiert und angepasst werden

muss, ist die Basiseinheit eines Wikis ein Artikel. Die gespeicherten Informationen können je nach Bedarf zusammengeführt oder geteilt werden. Ebenfalls ist der Informationskontext über die Verknüpfung der Seiten beliebig anpassbar. Der Mitarbeiter kann bei der Externalisierung seines Wissens auf bereits bestehende Inhalte oder auf noch nicht vorhandene Inhalte verweisen. Letzteres ermöglicht ihm, seinen Wissensbedarf implizit zu formulieren. Eine geeignete dieser Informationsbedarfe ermöglicht es anderen Nutzern, diese Lücken zu füllen.

Wikis eignen sich aufgrund des einfachen Prinzips der Verarbeitung und Verlinkung von Artikeln hervorragend, um gemeinschaftlich nicht hierarchische Wissenssammlungen zu erstellen und zu verwalten [MuDi06]. Der Wissensraum besitzt keine vordefinierte Struktur, wodurch das Informationsnetzwerk in einem kollaborativen, von den Mitarbeitern selbstgesteuerten Prozess entwickelt wird.

Wikis bieten häufig eine Übersicht, in welcher die aktuellen Artikeländerungen mit einem Zeitstempel auflistet sind. Da grundsätzlich alle Artikelversionen in der Historie gespeichert werden, können Änderungsumfang und -art mit einem Versionsvergleich ermittelt werden. Ein häufiges Problem der Wissensarbeit, dass zwar das Endprodukt wie beispielsweise ein Bericht sichtbar ist, aber nicht der Erzeugungsprozess, kann mit der Artikelhistorie gelöst werden. Wenn kollaborativ an einem Artikel über eine bestimmte Zeitspanne gearbeitet wird, werden die stetig vollzogenen Änderungen erfasst [EbGH06]. Der in den meisten Wikis vorhandene Beobachtungsmechanismus (*subscribe*) gestattet den Mitarbeitern einen bedarfsgerechten Umgang mit den verfügbaren Informationen, indem sie über Änderungen an abonnierten Seiten informiert werden.

Auch in Unternehmen lebt der Wiki-Ansatz, wie auch bei der Wikipedia, von der Eigeninitiative der Mitarbeiter. Daher sollte bei der Entscheidung für ein Wiki über entsprechende Anreizsysteme nachgedacht bzw. Maßnahmen ergriffen werden, die Nutzung im Unternehmen zu fördern. Beispielsweise mit Hilfe eines „Wiki-Champions“, einer Person, die den Wiki-Gedanken in einer Vorbildfunktion im Unternehmen vorantreibt. Wenn auch bestimmte Nutzer besonders wichtig sind, ist das Wiki ein Werkzeug, das die Bedeutung jedes Einzelnen anerkennt. Jeder einzelne Mitarbeiter wird als Kompetenzträger angesehen und befähigt, sein Wissen anderen frei zur Verfügung zu stellen. Die Objektivität und Qualität der Inhalte muss gesichert sein. Der Autor sollte eine Distanz zu den erzeugten

Inhalten aufbauen, damit eventuell durchgeführte Änderungen nicht auf sein Unverständnis stoßen.

Wissensgemeinschaften in Form von sozialen Netzwerken bilden sich in Wikis durch die gemeinsame Arbeit an Inhalten - den Artikeln. Unabhängig von der organisatorischen Verankerung finden sich Mitarbeiter aufgrund gemeinsamer Themeninteressen zusammen. Der Prozess der Netzwerkbildung erfolgt über die Suche der Mitarbeiter nach bestimmten Themen, die sie beispielsweise zur Problemlösung benötigen. Ist der gewünschte Inhalt vorhanden, erfolgt die Artikelnutzung in Form der Anwendung, Ergänzung oder Anpassung. Ein Wiki legt die Informationsherkunft offen, da die Autoren, die die gewünschten Inhalte eingestellt haben, identifizierbar sind und bei weiterführenden Fragen kontaktiert werden können. Auf Grundlage der Artikelbearbeitungen existiert eine Beziehung zwischen den Personen und diesen Artikeln, die eine Netzwerkbeschreibung erlauben. Das Netzwerk bildet sich sukzessive.

1.4 Selbstorganisation in sozialen Netzwerken

Allgemein betrachtet ist ein soziales Netzwerk eine eigenständige, zeitlich begrenzte Form der Koordination von Interaktionen, deren Kern eine vertrauensvolle und rücksichtsvolle Kooperation autonomer, aber interdependenter Akteure ist [Wey00]. Die sozialen Netzwerke besitzen Eigenschaften, die für fast alle Netzwerke gelten. Daher werden zunächst die grundlegenden Netzwerkeigenschaften vorgestellt, um dann im Speziellen auf die Wiki-Netzwerke (vgl. Abschnitt 3.4) einzugehen.

Viele Netzwerke besitzen Eigenschaften eines Kleine-Welt-Graphen (*small world*) geprägt [Wat99]. Diese besitzen eine geringe Dichte und einen hohen Clusterkoeffizienten. Eine geringe Dichte bedeutet, dass die Anzahl der vorhandenen Kanten gemessen an der Anzahl der möglichen Kanten gering ist. Trotz der relativ kleinen Kantenzahl ist der Durchmesser der Graphen (die größte Anzahl der Kanten, die mindestens durchlaufen werden müssen, um zwei beliebige Knoten zu verbinden) niedrig. Beispielsweise bedeutet das in einem Kommunikationsnetzwerk, dass eine Nachricht, die von einem Knotenpunkt zu dessen Nachbarn weitergegeben wird, in kürzester Zeit alle Knoten des Netzes erreicht hat. Dies ist insbesondere auf vereinzelt vorkommende Kanten zu besonders weit entfernten Knoten zurückzuführen. Kleine-Welt-Graphen bilden Cluster. Die Wahrscheinlichkeit, dass ein Knoten eine Verbindung zu einem benachbarten Knoten eingeht, ist sehr hoch. Die Wahrscheinlichkeit hingegen, dass er zu einem weiter entfernten Knoten verlinkt, ist

vergleichsweise gering. Der Clusterkoeffizient ist ein Maß für den Grad der Verlinkung in einem Netzwerk. Der lokale Clusterkoeffizient gibt an, wie viele Kanten ein Knoten im Verhältnis zur maximal möglichen Anzahl an Kanten besitzt.

Es wurden eine Vielzahl von Modellen entwickelt, um den Anforderungen der Small World Graphen zu genügen (z.B. [WaSt98], [Kle00], [BaAH99]). Barabási und Albert entwickelten ein Modell, in dem lediglich die Knoten spezifische Eigenschaften besitzen müssen [BaAH99]. In dem Modell weisen Knoten keine spezifische Anzahl von Kanten pro Knoten auf. Sie folgen einer logarithmischen Verteilung (*power law*). Es bedeutet, dass es sehr viele Knoten gibt, die relativ wenige Kanten haben. Andererseits gibt es aber auch einige Wenige besonders stark verlinkte Knoten in diesem Netz. Ein solcher, sehr gut vernetzter Knotenpunkt, ist eine Drehscheibe (*hub*). Des Weiteren haben Barabási und Albert das Wachstum solcher Netzwerke untersucht. Neue Knoten verbinden sich bevorzugt mit Knoten, die bereits viele Kanten besitzen (*preferential linking*). Small Networks sind robust und adaptiv bezüglich externer Störungen. Das bedeutet, dass die zufällige Entfernung von Knoten nur geringe Auswirkungen auf das Netzwerk hat, da die Mehrzahl der Knoten gering verbunden ist. Fallen dagegen hoch vernetzte Knoten aus, kann das Netzwerk stark geschwächt werden.

Netzwerke weisen emergente Eigenschaften auf, d.h. wenn ein kritischer Schwellenwert (*tipping point*) überschritten ist, agieren alle Knoten des Netzwerks als eine Einheit [Bar03]. Emergenz bedeutet dabei, dass autonomes Verhalten oder Eigenschaften auf einem höheren Ordnungsniveau (Makroebene) nicht direkt mit der unteren Ebene (Mikroebene) erklärbar sind [Ste99]¹. Der Transformationsprozess der Emergenz beruht auf bestimmten Regeln. Während des Transformationsprozesses treten neue Qualitäten auf, die nicht auf die Aufsummierung der Einzeleigenschaften zurückgeführt werden können. Die emergenten Eigenschaften eines Netzwerks basieren auf Selbstorganisation, einem Phasenübergang hin zu höherer räumlicher Ordnung (vgl. Abschnitt 3.1) [Hue06]. Unterschiedliche Disziplinen haben spezifische Selbstorganisationstheorien entwickelt [PaKn91], beispielsweise die biologische und kybernetische Systemtheorie von H.v. Förster, die Synergetik von H. Haken sowie die Chaostheorie von E. Lorenz und B. B. Mandelbrot.

¹ Es handelt sich um eine stark vereinfachte Definition, daher wird auf die Ausführung von Achim Stephan verwiesen, der die Merkmale emergentistischer Theorien ausführlich erläutert.

2 Forschungsfrage & gewähltes Vorgehen

Das Ziel dieses interdisziplinären Promotionsvorhabens ist, die Aspekte der Selbstorganisation, die mit der Nutzung von Wikis für das Wissensmanagement in virtuellen betrieblichen Wissensgemeinschaften verbunden sind, zu beschreiben, zu analysieren und ihre Ausprägung zu bewerten. Die Ergebnisse der Arbeit werden genutzt, um die Regeln zur Förderung von Selbstorganisationsprozessen in Unternehmen in ihren technischen (hinsichtlich der Wiki-Anwendung), kulturellen (hinsichtlich der sozialen Netzwerke) und organisatorischen (hinsichtlich der Wissensgemeinschaften) Gestaltungsbereiche abzuleiten.

Können Prozesse der Selbstorganisation in virtuellen Wissensgemeinschaften anhand der Netzwerkstrukturen von Wikis identifiziert werden? Um diese Fragestellung zu beantworten muss zunächst geprüft werden, ob Phasenübergänge innerhalb der Netzwerkstrukturen erkennbar sind (vgl. Abschnitt 3.1). Außerdem wird die Frage beantwortet, ob und wie Wikis erweitert werden können, um Selbstorganisation durch bestimmte Mechanismen zu fördern. Es wird aber nicht in die Wissensprozesse der Mitarbeiter eingegriffen, sondern deren „Spuren“ analysiert, um bestimmte Informationen zur Verfügung zu stellen.

Die Promotionsarbeit stellt zunächst den Stand der wissenschaftlichen Diskussion zur Selbstorganisation vor. Ausgewählte Erkenntnisse werden beschrieben und auf ihre Anwendbarkeit im betrachteten Forschungsfeld geprüft. Nach einer Einführung in die grundlegenden Ansätze im Wissensmanagement wird insbesondere auf die Rolle von virtuellen Wissensgemeinschaften eingegangen und diese einer kritischen Bewertung hinsichtlich der bestehenden informationstechnischen Unterstützung unterzogen. Danach folgt eine Einführung in die so genannte Soziale Software, welche als informationstechnische Hilfsmittel für das Wissensmanagement zunehmend Anwendung findet. Besonderes Augenmerk wird dabei auf die Nutzung von Wikis als Informationssammlung in virtuellen Wissensgemeinschaften gelegt. Die gewonnenen Erkenntnisse bilden den Ausgangspunkt für die Bestimmung von Kriterien für ein selbstorganisiertes Wissensmanagement. Das Vorgehen besteht im Kern aus drei Schritten.

Im ersten Schritt werden die innerhalb der Wiki-Software bestehenden Netzwerke beschrieben. Es konnten unter anderem bereits das Competence-Network (Verknüpfung Wiki-Autoren und Artikel über Artikelbeschreibungen), das Collaboration-Network (Verknüpfung von Autoren über Artikel-Bearbeitungen) und das Discussion-Network

(Verknüpfung von Diskussionsseiten über Beiträge von Autoren) identifiziert werden. Jedes einzelne Netzwerk wird zur Analyse bestimmter Phänomene genutzt. Die Ergebnisse der einzelnen Netzwerke werden zur Validierung miteinander verglichen.

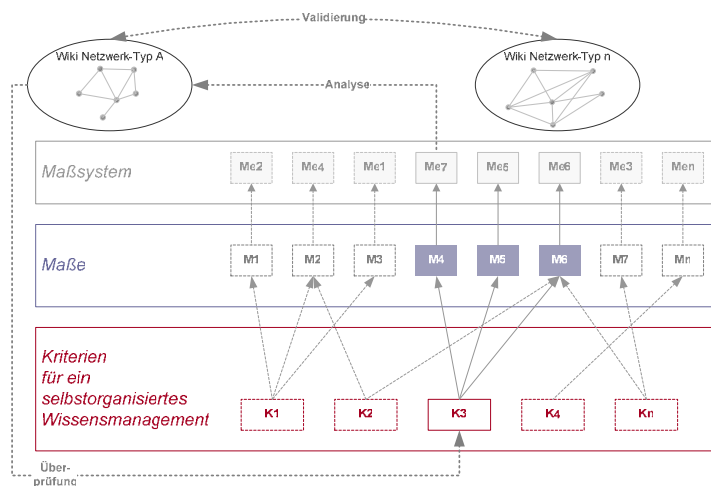


Abb. 1: Gewählte Vorgehensweise des Promotionsvorhabens

Im zweiten Schritt wird die Hauptthese der Arbeit durch Unterthesen operationalisiert. Die Unterthesen werden den Kriterien des Selbstorganisierten Wissensmanagements zugeordnet (vgl. Abb. 1). Für jedes Kriterium werden gewisse Merkmale bestimmt, die auf die Erfüllung des Kriteriums schließen lassen. Danach kann die Festlegung des zu verwendenden Netzwerktyps (vgl. Abschnitt 3.1) und eines entsprechenden Maßes (vgl. Abschnitt 3.2) erfolgen. Ein Test des Messsystems wird anhand von Wikipedia-Daten vorgenommen.

Der dritte Schritt umfasst die Anwendung dieses Messsystems auf reale Unternehmensanwendungen im Bereich Wissensmanagement. Es konnten bereits zwei Unternehmen gewonnen werden, die ihre Wiki-Daten für eine Analyse zur Verfügung stellen. Es handelt sich dabei um Unternehmen aus dem Bereich der Softwareentwicklung. Bei der Analyse kann so ein Teil der Software-Entwicklung sichtbar gemacht werden, der sonst verborgen ist. Das bestehende Messsystem wird auf seine Praktikabilität getestet, die definierten Maße in ihrer Aussagekraft geprüft und der Ansatz des selbstorganisierten Wissensmanagements untersucht.

Im letzten Schritt erfolgt die Evaluation der erhaltenen Analyseergebnisse. Zunächst werden die Software-Entwicklungsprozesse mit den Externalisierungsprozessen im Wiki abgeglichen. Danach werden selbstorganisierende Mechanismen in das Wiki integriert, wie

die Bewertung von Artikeln. Diese Erweiterungen dienen der gezielten Verankerung und Stärkung von Mechanismen der Selbstorganisation im Wiki. Letztlich erfolgt eine Abbildung der virtuellen sozialen Netzwerke auf die realen Netzwerke, um so die Software-Entwicklungsprozessen im Bedarfsfall umzugestalten.

Im Gegensatz zu den häufig angewandten linear verlaufenden Forschungsdesigns soll sich das hier angewandte Verfahren durch eine dynamischere Prozessualität und Zirkularität auszeichnen [Flic98]. Die einzelnen Phasen der Hypothesenformulierung, der Spezifikation des methodischen Vorgehens, der Datenerhebung und -analyse, der Hypothesenauswertung und der Ergebnisbeurteilung werden als Orientierungsrahmen verstanden, in welchem durchaus Rücksprünge vorgenommen werden. Die beschriebene Vorgehensweise wird im nächsten Abschnitt am Beispiel der Synergetik veranschaulicht.

3 Übertragung eines Selbstorganisationsansatzes auf das Untersuchungsgebiet

Dieser Abschnitt erklärt das gewählte Vorgehen anhand der Synergetik, die ein originär emergentistisches Problem verdeutlicht: die Entstehung neuartiger Systemeigenschaften [Ste99]. Das Wirkmodell der Synergetik wird auf das Untersuchungsgebiet der Wiki-Netzwerke übertragen. Es werden die ausgewählten Maßen der sozialen Netzwerkanalyse vorgestellt, mit welchen das Untersuchungsgebiet analysiert wird. Da die Interpretation der Maße kontextabhängig ist, werden ausgewählte Wiki-Netzwerke näher erläutert.

3.1 Grundlagen der Synergetik

In der Physik hat Hermann Haken das Forschungsgebiet der Synergetik („die Lehre vom Zusammenwirken“) gegründet. Die Synergetik ist als einzige Selbstorganisationstheorie in sich abgeschlossen und mathematisch formuliert [Hak90]. Den Anstoß zur Begründung der Synergetik gab die Entwicklung der Lasertheorie. Diese wird nun genutzt, um die zentralen Begriffe der Synergetik aufzuzeigen.

Ein Laser unterscheidet sich von einer üblichen Gasentladungsröhre lediglich durch die beiden Spiegel an den Endflächen der Röhre. Von außen wird eine Spannung angelegt, welche die Atome anregt. Durch die Veränderung der *Kontrollparameter*, hier in Form der Stromstärke, wird das System dynamisiert und von seinem Gleichgewichtszustand entfernt.

Der Gleichgewichtszustand ist dabei ein dynamisches Schwanken um einen stabilen Zustand [Bru02]. Die Art und der Grad der Ordnung des Systems werden über die *Ordnungsparameter*, hier Amplitude und Wellenlänge, bestimmt. Das Gesamtverhalten des Systems ist trotz der hohen Komplexität über diese Ordnungsparameter ausreichend beschreibbar.

Eine Lichtwelle entsteht, wenn ein Elektron von einem energiereicheren zu einem energieärmeren Zustand wechselt. Aufgrund der Spiegel verbleiben die Lichtwellen relativ lange im Laser und können mehr und mehr Elektronen anregen. In der Anfangsphase entstehen Lichtwellen unterschiedlicher Amplitude und Wellenlänge (Zunahme der *Fluktuationen*), die in Konkurrenz zueinander treten, um von den Elektronen Verstärkung zu erhalten. Das System befindet sich nun in der *Symmetriephase*, dem Zustand der maximalen Instabilität.

Die Elektronen verstärken die einzelnen Lichtwellen nicht in gleicher Weise. Sie geben ihre Energie mit geringem Vorrang an eine bestimmte Welle ab, die ihrem „inneren“ Rhythmus am nächsten kommt. Trotz dieser geringen Bevorzugung, wird diese Lichtwelle lawinenartig verstärkt und bricht die Symmetrie (*kritische Fluktuation*). In einem positiven *Rückkopplungsprozess* sammelt sich die Energie aller Elektronen in einer kohärenten Welle, die alle anderen unterdrückt. Die neu entstehende Welle bestimmt die Ordnung im Laser und übernimmt sozusagen die Rolle des *Ordnerns*. Durch den Ordner findet eine Komplexitätsreduzierung statt, da die einzelnen Elektronen zuzusagen *versklavt* werden [Hak88]. Eine Lasertätigkeit setzt bei Erreichen einer kritischen Anzahl von Laseratomen ein. Es kommt zu einem Wechsel von der Quantität zur Qualität.

Im folgenden Abschnitt wird das Wirkmodell der Synergetik auf die selbstorganisierenden Prozesse in einem Wiki übertragen.

3.2 Anwendung der synergetischen Prinzipien auf Wiki-Netzwerke

Die Synergetik ermöglicht die auf der Mikroebene verursachten Strukturbildungsvorgänge in einem Wiki zu untersuchen und die Voraussetzungen für die Bildung erfolgreicher makroskopischer Netzwerkstrukturen in einem Wiki zu bestimmen. Im Folgenden werden die Kernelemente der Synergetik auf das Untersuchungsgebiet übertragen.

Die Synergetik bietet den Rahmen für die Untersuchung der Strukturbildungsprozesse in Wiki-Netzwerken. Das Ziel ist es, abhängig von der Phase, in welcher sich die Wiki-Gemeinschaft befindet, die entsprechenden Maßnahmen abzuleiten, um den

Wissensaustausch und die -dokumentation zu fördern. Insbesondere die Übergangsphasen bei der Entstehung von neuer Ordnung (beim Laser das kohärente Licht) können genutzt werden, um die dynamischen Wissensprozesse in einem Wiki zu erklären. Dazu wird zunächst betrachtet, ob das untersuchte System eine Übereinstimmung der internen Systembedingungen zu den externen Rahmenbedingungen sowie ein Zusammenhang zwischen den Systemfaktoren Struktur und Verhalten aufweist.

Ein Wiki ist als offenes System in ständiger Interaktion mit seiner Umgebung. Dem Informationsraum werden fortwährend Informationen zugeführt und entnommen. Der Informationsraum wird in einzelne Artikel (Knoten) zergliedert, die über Wiki-Links (Kanten) miteinander verknüpft sind. Selbstorganisation im Sinne der Synergetik bezeichnet die Fähigkeit eines Netzes, bei der Veränderung der Kontrollparameter (beim Laser die Spannung) Übergänge zwischen verschiedenen Phasen aus seiner internen Dynamik heraus zu vollziehen. Im Sinne der Synergetik werden verschieden mögliche Anordnungen (Strukturen) der Komponenten als unterschiedliche Phasen bezeichnet. Die neuen Phasen sind dabei auch mit neuen makroskopischen Eigenschaften des Netzes verknüpft. Daher ist es erforderlich, diese Strukturen zu erkennen und zu analysieren, mit dem Ziel, positive emergente Effekte auszulösen und ungewollte Emergenz zu verhindern.

Das Mikrosystem eines Wikis bilden die einzelnen Informationsobjekte (externalisiertes Wissen) in Form von Artikeln. Es werden individuelle Entscheidungen für das Erstellen, Erweitern, Verändern etc. eines Artikels getroffen. Die Summe der individuellen Entscheidungen kann nun zu Veränderungen auf der makroskopischen Ebene, dem Wiki als Wissensmanagementwerkzeug, führen. Das Makrosystem ist die Summe der Knoten mit den jeweils definierten Beziehungen – das Netzwerk. Die strukturelle Einbindung des einzelnen Netzwerkakteurs hat Auswirkungen auf sein eigenes Verhalten sowie das Verhalten anderer, da der Informationsfluss zwischen den einzelnen Akteuren vom Verhalten aller beteiligten Akteure abhängig ist. Dies wird als Rückkoppelung bezeichnet.

Das System muss aus einer größeren Anzahl von Systemelementen bestehen, damit Ordnern (beim Laser die Lichtwelle beim Wiki ein Thema) entstehen können. Die konkrete Anzahl der benötigten Systemelemente hängt vom System und der Vernetztheit der Elemente ab. Es kann mit dem mathematischen Modell von Haken abgeschätzt werden [Hak90].

Der Ansatz wird nun an einem Beispiel verdeutlicht. Entschließen sich immer mehr Mitarbeiter, ihr Wissen im Wiki zu explizieren, wird auch eine zunehmende Anzahl an

Mitarbeitern das dort externalisierte Wissen nutzen. Dem gegenüber stehen der Effekt, dass die Anzahl Nutzer gemessen an der Mitarbeiterzahl zu gering ist und somit auch der Umfang des externalisierten Wissens. Daher werden auch nur wenige Mitarbeiter das System nutzen. In diesem Fall muss der Schwellenwert des Kontrollparameters „Autorenanzahl“ ermittelt werden, um korrigierend eingreifen zu können. Einen Anhaltspunkt für die Ermittlung des Schwellenwerts gibt die Wahrscheinlichkeitsverteilung nach Pareto [Par35]. Diese Verteilung beschreibt das statistische Phänomen, dass sich die Gesamtzahl der Wiki-Nutzer in etwa 20 Prozent Autoren und 80 Prozent Nicht-Autoren (Leser) aufteilt.

Die Veränderung eines Kontrollparameters bewirkt ab einem bestimmten Schwellenwert, dass sich ein System dynamisiert und seinen stabilen Gleichgewichtszustand verlässt. Durch die Zufallsschwankungen werden die Ordnungsparameter und damit das System destabilisiert. Schwankungen, die von außen auf das System wirken, werden als Störungen bezeichnet. Vom System selbst erzeugte Schwankungen werden Fluktuationen genannt [Hak88]. Eine Störung für die Selbstorganisation im Informationsraum ist beispielsweise eine Unternehmensführung, die sich entschließt, die Nutzung eines Wikis durch monetäre Anreize zu fördern. Eine Fluktuation tritt ein, wenn Autoren zunehmend demotiviert sind und von einer weiteren Nutzung dieses Wissensmanagementwerkzeuges absehen. Störungen und Fluktuationen sind normale Systemeigenschaften, die dazu führen, dass sich das System innerhalb eines Intervalls um einen stabilen Systemzustand bewegt. Es wird überprüft, wie Störungen zielgerichtet zur Förderung von Wissensaktivitäten eingesetzt werden können bzw. welche Störungen den Wissensaktivitäten entgegenstehen. Ein Ansatzpunkt ist die Installation eines „Wiki-Champions“ (siehe Abschnitt 1.3). In der hier vorgestellten Arbeit wird ermittelt, ab wann bestimmte Beziehungen von Kontrollparametern und Fluktuationen bestehen, die einen Phasenübergang in einen nicht gewünschten Effekt, beispielsweise Nicht-Nutzung des Systems bewirken.

Das kollektive Verhalten der Komponenten und der Ordnungsparameter bedingen sich gegenseitig. Es besteht eine zirkuläre Kausalität zwischen der Struktur und dem Verhalten des Netzwerkes – das Versklavungsprinzip (siehe Abschnitt 3.1).

Ordnungsparameter sind systembeschreibende Parameter. Sie entstehen aus der Interaktion der Systemelemente, wirken auf diese zurück und sind für die entstehende Ordnung charakteristisch [Hak88]. Bezogen auf die Wiki-Netzwerkstruktur sind mögliche

Ordnungsparameter die Gruppenkohäsion und die informelle Organisationsstruktur. Diese Ordnungsparameter müssen, um selbstorganisierende Prozesse innerhalb der Systemnutzung erkennen zu können, bekannt sein.

3.3 Ausgewählte Maße der sozialen Netzwerkanalyse

Die soziale Netzwerkanalyse (SNA) wird eingesetzt, um die Strukturbildungs- und Strukturänderungsvorgänge in einem Wiki zu untersuchen. Die SNA dient der systematischen Beschreibung und Analyse von sozialen Strukturen. Dabei wird die soziale Umgebung als Muster oder Regelmäßigkeit zwischen den interagierenden Akteuren interpretiert [WaFa97]. Die aufgefundenen Muster beschreiben das individuelle und kollektive Verhalten. Maße dienen der Analyse oder dem Vergleich von Netzwerken, von Elementen eines Netzwerks oder von Beziehungen eines Netzwerks. Zur Identifikation geeigneter Maße für die durchzuführenden Messungen wird das Goal-Question-Metric (GQM) Verfahren angewandt [BrRo88]. In diesem Verfahren muss zunächst das Ziel (goal) der Messung definiert, danach die zu beantwortende Frage (question) aufgestellt und das entsprechende Maß (metric) identifiziert werden.

Unter Nutzung des GQM-Verfahrens werden systematisch zu den bestehenden Hypothesen die entsprechenden Zielgrößen bestimmt. Aufbauend darauf werden die Maße bestimmt, die in der Lage sind, die spezielle Fragestellung zu beantworten. Die quantitativen Eigenschaften eines Netzwerks müssen in ihrem Kontext interpretiert werden, da die Bedeutung eines Maßes vom untersuchten Netzwerk abhängt. Daher werden in diesem Abschnitt zunächst die zwei gewählte Maße vorgestellt und im folgenden Abschnitt in den Kontext des Netzwerktyps gesetzt.

Die knotenbezogene Analyse zielt darauf ab, die einzelnen Knoten (Personen oder Artikel) eines Netzwerks auf ihre Eigenschaften und ihre Position im Gesamtnetzwerk zu betrachten. Es handelt sich um eine Untersuchung auf der Mikroebene des Systems. Insbesondere die Zentralitätsmaße können wertvolle Einblicke in die Strukturvorgänge des Wiki-Informationsraums geben. Grundsätzlich handelt es sich bei diesen Untersuchungen um Momentaufnahmen. Daher ist es notwendig, eine Datenreihe zu verschiedenen Zeitpunkten zu erheben. Die bekanntesten Zentralitätsmaße sind die Grad-, Nähe- und Betweenness-Zentralität [WaFa97]. Die *Informations-Zentralität* als weiteres Zentralitätsmaß ermöglicht die Analyse von Ausgleichsmechanismen im Wiki. Dieses Maß baut auf der *Relevanz* eines Akteurs auf. Die Relevanz eines Knotens bezeichnet die

Effektivität mit der dieser Knoten alle anderen Knoten im Netzwerk erreichen kann [Lat04]. Es wird die Fähigkeit eines Netzwerks bewertet, den Akteur, wenn er aus dem Netzwerk gelöst wird, zu ersetzen. Diese Fähigkeit wird durch die Differenz der Effizienz, mit der das Netzwerk vor und nach der Deaktivierung des Akteurs Informationen verbreiten kann, beschrieben. Die Informations-Zentralität berücksichtigt die Gewichtung von Verbindungen. Wege im Netzwerk werden dabei nicht nur nach ihrer Länge, sondern auch nach ihrem Gewicht verglichen. Das Gewicht beschreibt die Anzahl der Informationen, die über diese Verbindung ausgetauscht werden. Ein Akteur ist zentral bezüglich der Informations-Zentralität, wenn sein Ausfall die Effizienz des Netzwerks überdurchschnittlich beeinträchtigt, der Messwert also dementsprechend hoch ist [WaFa97]. Wiki-Nutzer mit einem hohen Messwert für die Informations-Zentralität sind essenziell für die Verteilung von Information innerhalb des betrachteten Netzwerks.

Bei der Untersuchung der makroskopischen Ebene hilft das Maß *Dichte* (vgl. 1.4). Sie gibt das Verhältnis von vorhandenen Kanten eines Netzwerks zur maximal möglichen Anzahl an. Die maximal mögliche Anzahl von Verbindungen wird erreicht, wenn jeder Knoten des Netzwerks mit jedem anderen Knoten verbunden ist, wodurch die Anzahl der Verbindungen durch die Anzahl der vorhandenen Knoten beschränkt ist. Existiert die maximale Anzahl an Verbindungen, hat das Netzwerk eine Dichte von eins und wird als vollständiges Netzwerk bezeichnet [WaFa97]. Ist der Messwert für die Dichte beispielsweise in einem Kommunikationsnetzwerk hoch, d.h. nahe an eins, so sind viele Informationskanäle innerhalb des Netzwerks ausgebildet. Andersherum sind Informationskanäle schlecht ausgebildet, wenn die Dichte des Nachrichtennetzwerks gering ist.

Bei der Analyse muss beachtet werden, dass Messwerte korrelieren. Daher können Maße nicht unabhängig voneinander betrachtet werden. Beispielsweise hat ein Netzwerk mit hoher Dichte eine geringe Informations-Zentralität, da viele Verbindungen existieren und Akteure dadurch leichter ersetzt werden können. Sinkt der Messwert für Dichte, steigt er gleichzeitig für Informations-Zentralität, da über weniger Verbindungen Informationen verteilt werden können.

3.4 Definition von Analysenetzenwerken in Wikis

Die quantitativen Eigenschaften eines Netzwerks müssen in ihrem Kontext betrachtet werden, da eine allgemeine, vom betrachteten Netzwerk unabhängige Interpretation nur schwer möglich ist. Die Kontextualisierung erfolgt über die Definition von Wiki-

spezifischen Netzwerken (vgl. Abschnitt 2). Analog zu den Maßen wird bei der Spezifikation der Netzwerke das GQM-Verfahren genutzt. Im Folgenden werden das *Wiki-Link-Network* und das *Information-Flow-Network* vorgestellt.

Artikel, die durch Wiki-Links miteinander verbunden sind, bilden das *Wiki-Link-Network*. Die Annahme ist, dass verlinkte Artikel thematisch miteinander in Beziehung stehen. Eine Gewichtung der Kanten kann über die Häufigkeit der Verlinkung erfolgen. Es handelt sich um ein gerichtetes Netzwerk, da die Verbindungen in Wikis nicht reziprok sind. Die Ermittlung des Clusterkoeffizienten (vgl. Abschnitt 1.4) im Zeitverlauf gibt Auskunft, wie sich die Themenstruktur ändert und befähigt die einzelnen Phasen der Themenstrukturierungen in einem Wiki zu bestimmen. Zu Beginn einer Wiki-Einführung werden die einzelnen Artikel noch voreinander losgelöst existieren. Erst bei Erreichen einer bestimmten Nutzerzahl (Kontrollparameter) werden strukturgebende Effekte eintreten und es kommt zum Entstehen einer Themenkohäsion.

Die Analyse des *Information-Flow-Network* identifiziert „interessante“ Artikel im Wiki. Als „interessant“ gelten Artikel, die häufig bearbeitet, geändert oder erweitert werden. Verbindungen zwischen Artikeln entstehen durch die Bearbeitung des gleichen Autors. Im Gegensatz zum *Wiki-Link-Network* bilden sich in diesem Netzwerk Zentren nicht durch inhaltliche Verknüpfungen, sondern durch Autoren, die Artikel bearbeiten. Es lassen sich Informationsflüsse innerhalb des Wikis analysieren, da sich in einer zeitlichen Betrachtung herausfinden lässt, welcher Artikel mit dem Wissenshintergrund des Inhalts anderer Artikel bearbeitet wurde. Des Weiteren erfolgt eine Untersuchung, bei welcher Veränderung der Kontroll- und Beschreibungsparameter in diesem Netzwerk Subnetzwerke entstehen. Subnetzwerke entstehen, wenn Autorengruppen losgelöst von anderen Autoren Artikel verfassen.

4 Aktueller Arbeitsstand & weiteres Vorgehen

Aktuell werden die einzelnen Ansätze im Bereich der Selbstorganisation erhoben und auf ihre Verwendbarkeit in dem Untersuchungsgebiet analysiert. Es wird insbesondere die Fragestellung behandelt, ob sich allgemeingültige Beschreibungskriterien für die Selbstorganisation ableiten lassen, die dann auf das Wiki-basierte Wissensmanagement übertragbar sind. Die identifizierten Maße der sozialen Netzwerkanalyse werden in den

verschiedenen Netzwerken interpretiert und anhand von Wikipedia-Daten auf ihre Aussagekraft geprüft. Nach erfolgreichem Abschluss werden beide Bereiche über die Merkmale zusammengeführt und in Unternehmen angewandt und die Ergebnisse evaluiert. Der geplante Abgabetermin der Dissertationsschrift ist Ende des Jahres 2007.

Literaturverzeichnis

- [BaAH99] Barabási, Albert-László, Albert, Reka, Hawoong, Jeong: Mean-field theory for scale-free random networks. *Physica A*, 272 (1999), S. 173-187.
- [Bar03] Barabási, Albert-László: *Linked*. Plume Printing, New York 2003.
- [BiMu06] Birn, Lukas, Müller, Claudia: Kollaboratives Dokumentieren mit Sozialer Software. *ERP-Management - Zeitschrift für unternehmensweite Anwendungssysteme* (2006) 2, S. 36-39.
- [BrRo88] Victor R. Basiliand, Rombach, H.Dieter: The TAME project: towards improvement-oriented software environments. *SoftwareEngineering, IEEE Transactionson*, 14 (1988) 6, S. 758–773.
- [Brun02] Brunken, R. *Verbesserte Anpassungsfähigkeit des Änderungsmanagement durch komplementäre Selbstorganisation*. Fraunhofer IRB Verlag, Stuttgart 2002.
- [Der99] McDermott, Richard: Nurturing Three Dimensional Communities of Practice: How to get the most out of human networks. *Knowledge Management Review*, Fall (1999), S.26- 29.
- [EbGH06] Ebersbach, A., Glaser, M., Heigl, R.: *Wiki- Web Collaboration*. Springer-Verlag Berlin Heidelberg, 2006.
- [Flic98] Flick, Uwe: *Qualitative Forschung*. Rowohlts Enzyklopädie, Reinbek 1998.
- [Hake88] Haken, H. *Erfolgsgeheimnisse der Natur. Synergetik: Lehre vom Zusammenwirken* Ullstein, Frankfurt/M. Berlin 1988.
- [Hake90] Haken, Hermann: *Synergetik: eine Einführung*. Springer-Verlag, Berlin 1990.

- [Hue06] Hütt, Marc-Thorsten: Was ist Selbstorganisaiton und was nützt sie zum Naturverständnis? In: Vec, Milos, Hütt, Marc-Thorsten, Freund, Alexandra M.: Selbstorganisation. Böhlau Verlag, Köln 2006.
- [Kle00] Kleinberg, Jon M.: The small-world phenomenon: An algorithm perspective. In: ACM Symposium on Theory of Computing (2000).
- [Klo06] Klobas, Jane: Wikis: Tools for Information Work and Collaboration. Chandos Publishing, Oxford 2006.
- [Lat04] Latora, Vito, Marchiori, Massimo: A measure of centralità based on the network efficiency. <http://www.citebase.org/abstract?id=oai:arXiv.org:cond-mat/0402050>, Abruf am 2006-11-29.
- [LeCu01] Leuf, Bo, Cunningham, Ward: The Wiki Way - Quick Collaboration on the Web. Addison-Wesley, New York 2001.
- [MaWY06] Majchrzak, A.; Wagner, C. & Yates, D. Corporate Wiki Users: Results of a Survey. <http://www.wikisym.org/ws2006/proceedings/p99.pdf> , Abruf am 2006-11-29.
- [MuDi06] Mueller, Claudia, Dibbern, Peter: Selbstorganisiertes Wissensmanagement in Unternehmen basierend auf der Wiki-Technologie - Ein Anwendungsfall. HMD – Praxis der Wirtschaftsinformatik, Heft 252 (2006) Dezember.
- [Nai82] Naisbitt, John: Megatrends. Heyne, München 1982.
- [Neu99] Neuweg, Hans-Georg: Könnerschaft und implizites Wissen. Zur lehr-lerntheoretischen Bedeutung der Erkenntnis- und Wissenstheorie Michael Polanyis. Waxmann Verlag, Münster et al. 1999.
- [NoRP00] North, Klaus, Romhardt, Kai, Probst, Gilbert: Wissensgemeinschaften – Keimzellen lebendigen Wissensmanagements. In: io-Management, Nr. 7/8 (2000), S. 52-62.
- [Par35] Pareto, Vilfredo: The mind and society. Harcourt Brace, New York 1935.
- [PaKn91] Paslack, Rainer & Knost, Peter: Zur Geschichte der Selbstorganisationsforschung. Kleine Verlag GmbH, Bielefeld 1990.

- [Rom02] Romhardt, K.: Wissensgemeinschaften. Orte des lebendigen Wissensmanagements. Versus Verlag Zürich, 2002.
- [Sco98] Scott, W. Richard: Organizations - Rational, Natural, and Open Systems. Prentice Hall, New Jersey 1998.
- [Ste99] Stephan, Achim: Emergenz - Von der Unvorhersagbarkeit zur Selbstorganisation. Dresden University Press, Dresden 1999.
- [TaCa93] Tapscott Don, Caston Art: Paradigm Shift: The new promise of information technology. McGraw-Hill Companies New York 1993.
- [WaFa97] Wasserman, Stanley, Faust, Katherine: Social network analysis: methods and applications. Cambridge Univ. Press Cambridge, 1997.
- [WaSt98] Watts, Duncan J., Strogatz, Steven H.: Collective dynamics of 'small-world' networks. Nature 393 (1998), S. 440-442.
- [Wat99] Watts, Duncan J.: Small Worlds. Princeton University Press, Princeton 1999.
- [Wen98] Wenger, Etienne: Communities of Practice: Learning, meaning, and identity. Cambridge University Press, Cambridge 1998.
- [Wey00] Weyer, Johannes: Soziale Netzwerke. Oldenbourg, München 2000.

Cannibalization Effects in Online & Offline Retailing

Consumers' Course

Gottfried Gruber

Institute for Management Information Systems
Vienna University of Economics and Business Administration
Augasse 2–6, 1090 Wien
gottfried.gruber@wu-wien.ac.at

Supervisor

Hans Robert Hansen

Institute for Management Information Systems
Vienna University of Economics and Business Administration
Augasse 2–6, 1090 Wien
hansen@wu-wien.ac.at

Abstract

The goal of this work is to define pricing strategies for online and offline retailing that minimize the effects of cannibalization between the different channels. I focus on homogeneous goods in a monopolistic market where the price can be set by the monopolist independently.

The thesis illustrates an alternative way of modelling a firm's profit function in an environment where the firm processes different sales channels. It shows, that there is no closed form solution possible within this framework and thus simulation is inevitable. From the simulation various suggestions can be drawn to support management decisions.

I will show how the number of consumers in the online channel influences the optimal pricing strategy for each channel. Furthermore, I will point out the effect of consumers, switching from the offline channel to the online channel, to the profit-maximizing pricing in each sales channel. Additionally sensitivity analysis of the profits on prices will be used to evaluate the impact of false estimates of the current market situation.

1 Introduction

Online marketing becomes more and more important. In some industries it has become a "must have" feature meanwhile. An additional online sales channel possibly reduces costs. Some authors argue that firms can add information to online offerings at low costs and thus achieve costs savings up to 25%, others state that firms conduct channel integration to save operational costs [AILW⁺97, AdBS04]. The unique features of the online channel allow firms to expand their offerings in a cheap way, which thereby help to serve consumers' needs better, and thus increase profits [AILW⁺97]. Furthermore, since menu costs are considerably low, firms are able to change prices more frequently and to a much finer extend than in their traditional channel. This price segmentation allows to skim consumer surplus better [LeGo02]. Furthermore, the online channel may lower transaction costs [Ward01, LiHu98].

Through the online channel new consumer groups can be reached. Recent literature asserts that an additional online channel increases market coverage and thereby firm's profit [FrFu03, BaLO⁺05]. Other studies highlight the demand expanding capacity of the online channel [BeTh04, GeGD02]. Positive effects are also attributed to the presumption, that the online channel strengthens relationships to existing consumers, and the firm receives a loyalty payoff from maintaining its online channel. Evidence from the travel industry and even from the outdoor industry show the profitableness of using the online medium to reinforce loyalty [ShSR03, WaGJ04].

Consumers may also obtain additional services by the online channel, thereby they may be more satisfied which in turn leads to increased loyalty. With the use of an adapted time allocation model it is possible to display that multi-channel retailers can indeed serve consumers needs better and thus hamper harmful switching loss [ReMc02]. Another study shows that loyalty can be leveraged if the firm is capable to put more weight on non-digital attributes of a product, which are more relevant for the subsequent purchase decision [LaSa99]. Evidence for loyal consumers being less price sensitive is also present in the coffee market [KrRa91].

Additionally there exist considerations that a firm may extract synergies from an additional online sales channel [BeTh04, AdBS04, StBA02]. Since each channel provides different advantages, consumers choose the optimal channel for their purchase. This increases consumer fit and for this reason the firm's profit [FoML04, Wiks05, ScGo02]. A spatial model shows that an information provider is able to increase profits by supplementing the direct channel by an alternative indirect channel [DeFS00]. Thus, consumer value can be enhanced by serving their needs better [BeTh04, AdBS04, SuTh04].

From the marketing perspective the online channel provides closer customer contact, which can be used for precise consumer profiling [TsAg04, LeGo02].

The Internet is often described as a "friction free" market [Bako01]. Increased competition should result in lower prices and less price dispersion. Furthermore, the distinctive features of the Internet, most notably reduced search costs, should increase price sensitivity. All together the friction free market should increase efficiency such that total welfare is increased [Bako01, ALLW⁺97, BaLO⁺05].

But despite these advantages many practitioners believe in a threat from the friction free market. Indeed reduced search costs encourage price competition, which firms try to prevent by introducing barriers [ALLW⁺97, Odly96, Bako97, Salo79].

Thus, the fear of cannibalization seems to be well grounded. Cannibalization determines sales shift from the entrenched channel to the new online channel, which may not increase profits rather decrease them [ALLW⁺97, BIWi89, MeMa01, SrRG05, StBA02, HaMa06]. Further, cannibalization will be more likely, the higher the perceived similarity between products [HaKe79]. Especially the Internet sales channel offering digital information goods is exposed to cannibalization. Some authors state that issues like cannibalization, channel coordination and channel conflicts may be more pronounced due to the nature of the Internet [Bala98]. Even worse, the online channel may not necessarily enhance consumer spending, partly due to a lack of cross selling potential [SuTh04].

The newspaper and the music CDs industry both indicate signs of cannibalization, and it was confirmed that cannibalization may increase as the Internet becomes more mature [DeGG⁺02, BiNa03]. Additional studies reveal figures of cannibalization and decreasing returns on consumer durable and apparel products [AnMN05]. Especially homogeneous goods, in particular digital information goods force firms to compete fiercely on prices because they contain no other differentiation feature. The unique features of digital information goods make particularly these firms very exposed to cannibalization [Bail98, ShVa99].

However, the online sales channel is not only a threat it may also be a chance. Some studies even argue that there exists no cannibalization effect since online search generates offline sales [Ward01]. But if the online sales channel becomes attractive, pricing strategies for this channel should also become relevant.

Therefore a proper pricing strategy is vital for firms doing business through different sales channels with homogeneous goods, since the price is closely related to the profit of a firm. A wrong pricing policy could harm firm's profit or even drive it to bankruptcy. For sure, wrong pricing causes deadweight loss. Deadweight loss denominates loss which could have been avoided by a proper allocation of resources, i.e. in this case an optimal pricing in both sales channels. First, deadweight loss occurs because a price too high lets some consumers

forgo a purchase which would otherwise have taken place and thus lowers a firm's profit. Second, a price too low would bring many consumers, but *ceteris paribus* due to low prices the profit would not be adequate. Therefore optimal pricing strategies are the passport to maximize profit in industries, which sell homogeneous products on the Internet as well as through traditional channels [Vari95, ShVa99, Lieb02, Skie00]. In general the price is always an important competitive issue in satisfying consumers [WaGJ04].

I want to support firms in a comparable situation with normative guidance for their pricing decisions. Market studies have some limitations. Inevitably they could harm a firm's business. In reality it causes danger to play around with prices to measure different effects. Thus, I propose a model to quantify cannibalization and to assist firms with pricing suggestions in a laboratory like environment.

The aim of this thesis is to develop a simulation tool for multichannel retailers to obtain computer-assisted optimal pricing strategies, which maximize profits. Starting with a simple market model suggestions for the management should be deduced. Stochastic simulations will be used to search for profit maximizing prices at different market conditions. In a *pari passu* manner decisive parameters will be changed to cover all reasonable market conditions and derive optimal pricing behavior, and the impact of deviating from profit maximizing pricing.

I will show how the number of consumers in the online channel influences the optimal pricing strategy for each channel. Furthermore, the effect of switching consumers in conjunction with the number of total available consumers to the profit-maximizing pricing in each sales channel will be highlighted. I will provide guidance if the demand shape of the channel is not perfectly known, and I will show the influence of a changing demand shape on the optimal pricing strategy.

2 Literature Review

Pricing is always an important issue. Especially pricing in the Internet receives growing attention as this market promises higher returns. It may also facilitate benefits for both sellers and buyers [Simo92, HaHa99, Bran01]. Numerous authors are concerned about the price changes required for doing business on the Internet sales channel. Many studies investigate the issue of the price level, the price elasticity and the price dispersion. These factors define a price structure in a market as a whole. I want to give a short overview on relevant research results on each mentioned issue.

2.1 The case of price levels

Against intuition prices in the online channel may be higher. Online prices may be higher due to differentiation, which increases equilibrium prices [Kuks04]. In eBay auctions trusted sellers can obtain a premium compared to sellers with many negative ratings [BaPa02]. Above there exists also empirical evidence of higher prices in the online channel, indicating consumers' willingness to pay a premium for convenience [Bail98, Ho-G98].

An alternative interpretation may be given concerning the maturity of the Internet. If the reach of the Internet is small we will observe high prices, but as soon as the Internet becomes more and more mature online prices will fall. However, there will be no discount to offline prices in a world with a mature Internet [Zett00]. In the same direction goes the argument of online well adopted products having lower prices, but not well adopted products display higher prices in the online channel [Bala98].

On the opposite, increased price competition should, by standard economic argument, drive prices down [BrKa00]. Numerous studies show that prices are distinctive lower in the online channel than in the offline channel. Especially digital information goods are found to have lower prices in the online channel [PaRS02, BrSm00, LeGo02, AnSh04]. Even in the car retailing industry the Internet is capable to lower prices for new cars [ZeMS06].

Ambiguous results on the price level in the online channel may indicate, that price levels are determined by the maturity of the Internet, the adoption of a certain product for selling through the Internet, the information strategy of the firms and the competitiveness of the market.

2.2 The case of price elasticity

Price elasticity describes consumers' reactions on marginal price changes. Against common knowledge, online consumers seem to be less price sensitive than offline consumers [DeRW98]. A number of studies found that improvements in quality and service will lower online consumers' price sensitivity even below offline consumers' price sensitivity [AllW⁺97, ShRP99, LyAr00].

Contrary, transparency in the online sales channel would cause consumers to be more price sensitive in the online channel than in the offline channel. Theoretical studies as well as empirical studies on tax rates and groceries all claim increased price sensitivity of online consumers [Gool00, BuHK⁺92, DaWD03].

It seems that price elasticity is strongly determined by the product class. Homogeneous products will show higher price elasticity compared to products which can be differentiated by their features. This asymmetry may also be an outcome of different search frictions.

2.3 The case of price dispersion

Price dispersion is an indicator of the competitiveness of a market. Higher price dispersion in the online channel than in the offline channel indicates that firms can avoid price competition by differentiating themselves with quality or service. Existing literature proves evidence for substantial price dispersion in the online channel for travelling agents, retailers in the books and CDs market and other products [ClHH98, Bail98, BrSm00, AnSh04]. Even theoretical models yield comparable results [IyPa03, BaMo01].

On the opposite are findings of lower price dispersion online, which imply higher price competition in this sales channel [Wern94, MoZS01]. Less price dispersion in the online channel than in the offline channel is found by studies on the car retail industry, CDs, DVDs, hardware, software and consumer electronics [MoZS01, FaXi01, PaSR02]. Even life insurances display less price dispersion due to online price comparison sites [BrGo02]. Overall there seems to exist no clear assertion concerning price dispersion, because some authors argue that price dispersion depends on the number of firms filling a certain market [BaMS04].

It might be, that price dispersion is a function of the product class, the number of firms in that market and the competitiveness of that market, as well as the brand strength of the incumbent firms.

2.4 The case of search costs

Search costs became an economic topic since the seminal work of Stigler [Stig61]. For the Internet it emerges also an important topic because search costs are assumed to decrease with the adoption of the Internet. The Internet encourages consumers to undertake unimpeded search across stores [AILW⁺97]. Reduced search costs result in increased competition and thereby in reduced prices [Bako01, Bako97]. Further, increased competition makes it harder for firms to generate profits [Lieb02]. Because of this, firms use brands to increase search costs and prevent price competition [BeDS96].

However, lower search costs also allow firms to better monitor their competitors. This may foster collusion which increase firms' revenues [CaRM05]. Further, firms may provide better consumer fit, since lower search costs help firms to identify qualified consumers. The Internet enables even profiling and monitoring back such strategies [Bako01, LeGo02]. Nevertheless online search costs may be not trivial [LyAr00]. Even if prices can easily be found, perceived search costs may be significant. Search costs of zero would imply unreasonable consideration sets for consumers, i.e. consumers consideration sets may be overestimated frequently [MeRS03]. Further there may exist an asymmetric search behavior. If search costs are lowest consumers tend to search too less and vice versa

[ZwRL⁺03]. Thus, consumers indeed do not always search for lowest price [SmBr01]. Empirical evidence highlights this phenomenon. Results show households are visiting on average 1.2 book sites and 1.3 CD sites prior to their purchase decision [JoMF⁺04]. Such behavior explains increased prices and profits for firms, if search costs become relevant, especially for low price products like books and CDs [LaSa99].

3 Theoretical Model

The basic model in this study consists of one monopolist offering a single homogeneous product through an online and an offline sales channel. Consumers are assumed independent in each sales channel and offline consumers are allowed to switch to the online channel. The number of maximal obtainable consumers in each sales channel is limited above and consumers buy one unit of the product merely. Consumers are also not fully rational, thus exhibiting inertia. Further no costs and an exogenous given switching fraction $s \in [0, 1]$ are assumed. The prices in each channel are given by $p_B, p_O \in [0, 1]$ ¹. Consumers reservation prices are distributed by the cumulative probability functions F_B and F_O .

In the following paragraph I give a shot summary of the results from previous work. First, profit may be increased by introducing an online sales channel. This is due to its independent demand and its function as "collecting tank" for switching consumers. Second, it pays to increase loyalty. If (channel) loyalty is high, consumers from the traditional sales channel are less willing to switch to the online channel and as the simulations show, a lower switching probability results in higher sales. The firm can define differentiated prices in each channel irrespectively of the consumer base and perfectly skim consumer surplus almost like a monopolist. Third, as the switching probability increases and the online consumer base becomes larger the firm should equalize the online and the offline price. The larger the online consumer base the more the offline price should be reduced and the less the online price should be increased to meet each other. Furthermore, the higher the switching probability the faster the firm should equalize both prices [Grub06]. Following the results from that first work further investigations on the sensitivity of the results according to the demand structure are necessary. A sensitivity analysis of prices should give some insights on the necessary accuracy of prices in both sales channels to achieve a certain amount of the maximum obtainable profit regarding to a specific market environment. Another interesting question refers to the profit loss if the firm recognizes the current market environment correctly, i.e. accurate estimates of the switching probability

¹B denotes the offline and O the online price. Further on: B denotes offline related variables and O online related parameter

and the online consumer base, but instead of commanding optimal prices it just strikes the price within an 95% confidence interval. The impact of an incorrect market valuation is another important issue. What happens to the profit of the firm if it has predefined pricing strategies for particular market environments, but fails to recognize the current market environment, i.e. false estimates of the switching probability and the online consumer base, and therefore applies an inappropriate pricing scenario. Things may even be worse, if the firm has no clear pricing scenarios at hand and fails to identify the current market environment.

The impact of those failures should be explained and proper defensive strategies should be suggested upon findings from simulations.

4 Extended Model

To bring more reality in the basic model I introduce different fixed \hat{c} ² and variable costs c for each channel [Grub06]. Additionally I use a stochastic switching probability F_S because it seems plausible that this parameter is not measurable easily. Further I endogenize the switching probability F_S . I assume that the probability of switching is implicitly given by both probability functions of the reservation prices. For example, if search costs are high I would expect the offline reservation prices to have a positive kurtosis, i.e. being flatter and the online reservation prices to be "relatively" more skewed to the left, i.e. concentrated around lower prices compared to the offline reservation prices. High search costs prohibit searching for the lowest price, hence consumers buy where they are. In the offline sales channel the probability function of the reservation prices will converge to an uniform distribution, or exactly uniform if the consumer has to take the first offer, or leave it in case [Stig61]. In this situation consumers can not learn from additional price quotes in the offline channel. Since per definition, obtaining information is cheaper in the online channel, it is evident that the density function of the reservation prices of the online consumers should be more skewed to the left compared to the density function of the offline consumers [Bako97, LaSa99]. But the flat shape of both density functions imply low switching probabilities, which I would expect in case of high search costs. From this it seems, that any combination of the distribution of the reservation prices in both channels will implicitly give the corresponding switching probability for that given product in that market. I assume that the amount of switching is determined by the difference of the online and the offline price. If $p_B - p_O = 0$ then there occurs no switching and even if, it should have no effect on the profit since prices are equal, otherwise if the difference is

²The dot marks a placeholder for either B or O

one all offline consumers switch to the online channel. From the definition of p_B and p_O follows, that the difference is also in between zero and one, i.e. $p_B - p_O \in [0, 1]$

Lets define a cumulative probability function F_S of the not-accepted price differences between both sales channels. Thus, if the price difference is zero, all consumers accept that difference, and hence no consumer intends to switch to the online channel, in other words the switching probability is zero. So the probability of the not-accepted price differences is a perfect proxy for the switching probability. Lets define the not-accepted price difference probability as a function of the probability functions of the reservation prices of both the offline and the online consumers $F_S = f(F_B, F_O)$. Thus the fraction of consumers not accepting the price difference between both sales channels and therefore switching to the online channel, with f_S as the not-accepted price difference density, is given as

$$s(F_B, F_O) = \int_0^{p_B - p_O} f_S(x) dx \quad (1)$$

To make it easier to read, I will refer to the result of Equation 1 further on as switching probability. Plugging this result into the formulas from the basic model (see [Grub06]) I get the number of buyers in each channel by

$$n_B = \widehat{b}_B \left(1 - \int_0^{p_B} f_B(x) dx \right) (1 - s(F_B, F_O)) \quad (2)$$

$$n_O = \widehat{b}_O \left(1 - \int_0^{p_O} f_O(x) dx \right) + \widehat{b}_B \left(1 - \int_0^{p_B} f_B(x) dx \right) s(F_B, F_O) \quad (3)$$

The profit function and the optimizing task is pretty the same as in the basic model [Grub06]. Since costs are introduced I get the following profit function, where c_B and c_O are the corresponding variable costs and \widehat{c}_B and \widehat{c}_O the appropriate fixed costs.

$$\pi(p_B, p_O) = (p_B - c_B) n_B + (p_O - c_O) n_O - (\widehat{c}_B + \widehat{c}_O) \quad (4)$$

Without loss of generality I can assume the fixed costs being zero, which is valid from now on. There is also no closed form solution available and so I conduct another simulation. Simulations are not common in marketing science, nevertheless successfully done [StTe02]

5 Further Research

5.1 Advanced Model

In this part I extend the "Extended Model" from chapter 4 for the multi-product firm by setting up different probability functions of reservation prices for various products. Figure 1 depicts the case of two products. For each product there exists a maximum

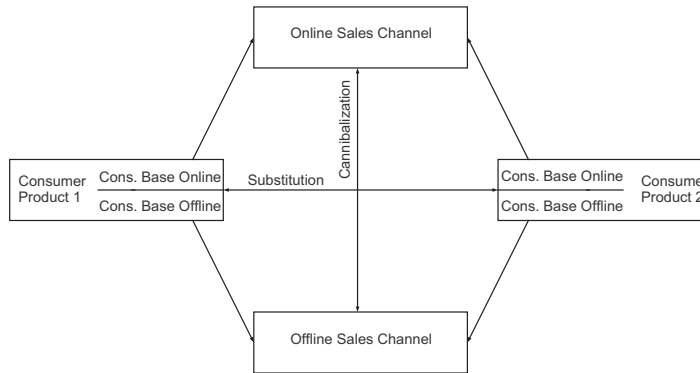


Figure 1: The Way of Consumers

obtainable consumer base for the online and the offline channel. As in the prior models, a consumer buying, say product 1 can switch from the offline channel of this product to the online channel of the same product and, in this model vice versa. Furthermore, consumers not buying product 1 because of their lower reservation price may consider purchasing product 2 alternatively. This effect is related to the substitution effect or the cross-price elasticity of demand [MaWG95].

This model can also be useful to provide insights in two firms competing with one homogeneous product in the same market, and both firms providing an offline as well as an online sales channel.

5.2 Sensitivity Experiment

To get deeper insights in the price sensitivity I will undertake an experiment to test at which price level consumers switch from their desired distribution channel to an alternative.

The setup is as follows. Test persons are confronted with a simple web frontend which asks them to purchase a certain product from one of the listed stores. The stores do not have known brands, but some are described such that the test person will associate the description with their best known brand. Other stores are described in terms of payment,

duration of delivery and distance from the test persons home or working place. In the first round all prices across all stores are equal, thus it is ensured that the test person searches for its preferred shop to undertake the purchase. After that purchase the price of the item in the selected shop will be increased for a certain amount, whereas all other prices stay the same. This procedure will be continued as long as the test person sticks to its initial choice. Afterwards the price of the "second best" shop will be increased until the next switching occurs. This will be continued until the third switching occurs. Afterwards the whole procedure starts again with another product. This is an aggregate view of the world, but it should give insights in the interconnection between price level and the likelihood of switching. These results should influence the endogenization process of the switching probability. I expect to receive an empirical distribution of accepted price differences, which is indeed an excellent statistical description of the switching probability.

5.3 Agent Based Simulation

In a further research direction I will model consumers as autonomous agents. Each consumer has its unique setup related to the finding in the stochastic simulations, and acts according to this rules [KeSB⁺00]. With that technique the course of each consumer can be tracked exactly and the driving parameter can be identified. One can see what action a specific consumer undertakes, and one can also investigate the parameter which influenced the consumer to perform that specific action.

This can be achieved with a landscape model for an agent-based simulation [EpAx96, Tern01, FlPU01]. The goal is to model the demand for a homogeneous good in a monopolistic market with two distribution channels. There are two independent "price planes" on which the prices are distributed uniformly. The price planes just reflect the firms probability of picking a price within zero and one at random.

Consumers are modeled as autonomous agents who can choose their desired price within the price planes. Each consumer has its own setup according to price sensitivity, information and search costs, which influences the switching probability, if it is possible for the agent to switch, and which determines the "loyalty" to a specific price.

In the calibration step the firm operates in the offline channel only. The firm can choose any price in the offline channel and then the consumer "sentiment" is adjusted to mimic the current demand shape, i.e. the consumers setup needs to be calibrated.

Afterwards the online channel is introduced. Now consumers of the firm are able to switch to the online channel and the firm can achieve additional consumers in the online channel, where the firm can also set its price independently. The question is how to set prices in the different channels to maximize profit and what parameters drive consumers to different channel choices.

References

- [AdBS04] *Adelaar, Thomas; Bouwman, Harry; Steinfeld, Charles*: Enhancing customer value through click-and-mortar e-commerce: Implications for geographical market reach and customer type. In: *Telematics and Informatics* 21 (2004) 2, S. 167–182.
- [AILW⁺97] *Alba, Joseph; Lynch, John; Weitz, Barton et al.*: Interactive home shopping: Consumer, retailer, and manufacturer incentives to participate in electronic marketplaces. In: *Journal of Marketing* 61 (1997) 3, S. 38–53.
- [AnMN05] *Ansari, Asim; Mela, Carl F.; Neslin, Scott A.*: Customer Channel Migration. Columbia University School of Business, NY.
- [AnSh04] *Ancarani, Fabio; Shankar, Venkatesh*: Price Levels and Price Dispersion Within and Across Multiple Retailer Types: Further Evidence and Extension. In: *Journal of the Academy of Marketing Science* 32 (2004) 2, S. 176–187.
- [Bail98] *Bailey, Joseph P.*: Intermediation and Electronic Markets: Aggregation and Pricing in Internet Commerce. Dissertation, Massachusetts Institute of Technology, 1998.
- [Bako97] *Bakos, Yannis J.*: Reducing buyer search costs: Implications for electronic marketplaces. In: *Management Science* 43 (1997) 12, S. 1676–1692.
- [Bako01] *Bakos, Yannis J.*: The Emerging Landscape for Retail E-Commerce. In: *The Journal of Economic Perspectives* 15 (2001) 1, S. 69–80.
- [Bala98] *Balasubramanian, Sridhar*: Mail versus Mall: A Strategic Analysis of Competition between Direct Marketers and Conventional Retailers. In: *Marketing Science* 17 (1998) 3, S. 181–195.
- [BaLO⁺05] *Bakos, Yannis J.; Lucas Jr., Henry C.; Oh, Wonseok et al.*: The Impact of E-Commerce on Competition in the Retail Brokerage Industry. In: *Information Systems Research* 16 (2005) 4, S. 352–374.
- [BaMo01] *Baye, Michael R.; Morgan, John*: Information Gatekeepers on the Internet and the Competitiveness of Homogeneous Product Markets. In: *American Economic Review* 91 (2001) 3, S. 454–474.
- [BaMS04] *Baye, Michael R.; Morgan, John; Scholten, Patrick*: Price Dispersion In The Small And In The Large: Evidence From An Internet Price Comparison Site. In: *Journal of Industrial Economics* 52 (December 2004) 4, S. 463–496.

- [BaPa02] *Ba, Sulin; Pavlou, Paul A.*: Evidence of the Effect of Trust Building Technology in Electronic markets: Price Premiums and Buyer Behavior. In: MIS Quarterly 26 (2002) 3, S. 243–268.
- [BeDS96] *Bergen, Mark; Dutta, Shantanu; Shugan, Steven M.*: Branded Variants: A Retail Perspective. In: Journal of Marketing Research 33 (1996) 1, S. 9–19.
- [BeTh04] *Berman, Barry; Thelen, Shawn*: A guide to developing and managing a well-integrated multi-channel retail strategy. In: International Journal of Retail and Distribution Management 32 (2004) 3, S. 147–156.
- [BiNa03] *Biyalogorsky, Eyal; Naik, Prasad*: Clicks and Mortar: The Effect of On-line Activities on Off-line Sales. In: Marketing Letters 14 (2003), S. 21–32.
- [BlWi89] *Blattberg, Robert C.; Wisniewski, Kenneth J.*: Price-Induced Patterns of Competition. In: Marketing Science 8 (1989) 4, S. 291–309.
- [Bran01] *Brandtweiner, Roman*: Report Internet-Pricing: Methoden der Preisfindung in elektronischen Märkten. 1. Aufl., Symposium Publishing, Düsseldorf 2001.
- [BrGo02] *Brown, Jeffrey R.; Goolsbee, Austan*: Does the Internet make markets more competitive? Evidence from the life insurance industry. In: The Journal of Political Economy 110 (2002) 3, S. 481–507.
- [BrKa00] *Brynjolfsson, Erik; Kahin, Brian (Hrsg.)*: Understanding the Digital Economy: Data, Tools, and Research. MIT-Press, Cambridge 2000.
- [BrSm00] *Brynjolfsson, Erik; Smith, Michael D.*: Frictionless commerce? A comparison of Internet and conventional retailers. In: Management Science 46 (2000) 4, S. 563–585.
- [BuHK⁺92] *Burke, Raymond R.; Harlam, Bari A.; Kahn, Barbara E. et al.*: Comparing Dynamic Consumer Choice in Real and Computer-Simulated Environments. In: The Journal of Consumer Research 19 (1992) 1, S. 71–82.
- [CaRM05] *Campbell, Colin; Ray, Gautam; Muhanna, Waleed A.*: Search and Collusion in Electronic Markets. In: Management Science 51 (2005) 3, S. 497–507.
- [CIHH98] *Clemons, Eric K.; Hann, Il-Horn; Hitt, Lorin M.*: The Nature of Competition in Electronic Markets: An Empirical Investigation of Online Travel Agent Offerings. The Wharton School of the University of Pennsylvania.

- [DaWD03] *Danaher, Peter J.; Wilson, Isaac W.; Davis, Robert A.*: A Comparison of Online and Offline Consumer Brand Loyalty. In: *Marketing Science* 22 (2003) 4, S. 461–476.
- [DeFS00] *Dewan, Rajiv; Freimer, Marshall; Seidmann, Abraham*: Organizing Distribution Channels for Information Goods on the Internet. In: *Management Science* 46 (2000) 4, S. 483–495.
- [DeGG⁺02] *Deleersnyder, Barbara; Geyskens, Inge; Gielens, Katrijn et al.*: How cannibalistic is the Internet channel? A study of the newspaper industry in the United Kingdom and The Netherlands. In: *International Journal of Research in Marketing* 19 (2002), S. 337–348.
- [DeRW98] *Degeratu, Alexandru; Rangaswamy, Arvind; Wu, Jeremy*: Consumer Choice Behavior in Online and Regular Stores: The Effects of Brand Name, Price, and Other Search Attributes. In: *Collected Working Papers, Marketing Science and the Internet*. MIT Sloan School 1998.
- [EpAx96] *Epstein, J.M.; Axtell, R.*: *Growing Artificial Societies: Social Science from the Bottom Up*. MIT-Press, Cambridge 1996.
- [FaXi01] *Fang-Fang; Xing, Xiaolin*: Will the growth of multi-channel retailing diminish the pricing efficiency of the Web? In: *Journal of Retailing* 77 (2001) 3, S. 319–333.
- [FIPU01] *Flentge, Felix; Polani, Daniel; Uthmann, Thomas*: Modelling the Emergence of Possession Norms using Memes. In: *Journal of Artificial Societies and Social Simulation* 4 (2001) 4.
- [FoML04] *Fox, Edward J.; Montgomery, Alan L.; Lodish, Leonard M.*: Consumer Shopping and Spending across Retail Formats. In: *Journal of Business* 77 (2004) 2, S. 25–60.
- [FrFu03] *Friedman, Lawrence G.; Furey, Timothy R.*: *The Channel Advantage*. Butterworth-Heinemann, Oxford 2003.
- [GeGD02] *Geyskens, Inge; Gielens, Katrijn; Dekimpe, Marnik G.*: The Market Valuation of Internet Channel Additions. In: *Journal of Marketing* 66 (2002) 2, S. 102–119.
- [Gool00] *Goolsbee, Austan*: In A World Without Borders: The Impact Of Taxes On Internet Commerce. In: *The Quarterly Journal of Economics* 115 (2000) 2, S. 561–576.

- [Grub06] *Gruber, Gottfried*: Cannibalization Effects in Online & Offline Retailing. In: 2006 Networking and Electronic Commerce Research Conference, NAEC 2006 2006 S. 57–70.
- [HaHa99] *Hanson, Ward A.; Hansson, Ward*: Principles of Internet Marketing. South-Western College Publishing Company 1999.
- [HaKe79] *Harvey, Michael G.; Kerin, Roger A.*: Diagnosis and Management of the Product Cannibalism Syndrome. In: University of Michigan Business Review 31 (1979) 6, S. 18–29.
- [HaMa06] *Hansen, Hans R.; Madlberger, Maria*: Beziehungen zwischen dem Internet-Vertrieb und anderen Absatzwegen im Einzelhandel. In: *Wirtz, B. W.* (Hrsg.): Handbuch Multi-Channel-Marketing. Gabler, Wiesbaden 2006, (forthcoming).
- [Ho-G98] *Ho-Guen, Lee*: Do electronic marketplaces lower the price of goods? In: Communications of the ACM 41 (1998) 1, S. 73–80.
- [IyPa03] *Iyer, Ganesh; Pazgal, Amit*: Internet shopping agents: Virtual co-location and competition. In: Marketing Science 22 (2003) 1, S. 85–106.
- [JoMF⁺04] *Johnson, Eric J.; Moe, Wendy W.; Fader, Peter S. et al.*: On the Depth and Dynamics of Online Search Behavior. In: Management Science 50 (2004) 3, S. 299–308.
- [KeSB⁺00] *Kearney, P.; Smith, R.E.; Bonacina, C. et al.*: Integration of Computational Models Inspired by Economics and Genetics. In: BT Technology Journal 18 (2000) 4, S. 150–161.
- [KrRa91] *Krishnamurthi, Lakshman; Raj, S.P.*: An Empirical Analysis of the Relationship Between Brand Loyalty and Consumer Price Elasticity. In: Marketing Science 10 (1991) 2, S. 172–183.
- [Kuks04] *Kuksov, Dmitri*: Buyer Search Costs and Endogenous Product Design. In: Marketing Science 23 (2004) 4, S. 490–499.
- [LaSa99] *Lal, Rajiv; Sarvary, Miklos*: When and How is the Internet Likely to Decrease Price Competition? In: Marketing Science 18 (1999) 4, S. 485–503.
- [LeGo02] *Lee, Zoonky; Gosain, Sanjay*: A longitudinal price comparison for music CDs in electronic and brick-and-mortar markets: Pricing strategies in emergent electronic commerce. In: Journal of Business Strategies 19 (2002) 1, S. 55–71.

- [Lieb02] *Liebowitz, Stan*: Re-Thinking the Network Economy: The True Forces That Drive the Digital Marketplace. American Management Association 2002.
- [LiHu98] *Liang, Ting-Peng; Huang, Jin-Shiang*: An empirical study on consumer acceptance of products in electronic markets: A transaction cost model. In: *Decision Support Systems* 24 (1998) 1, S. 29–43.
- [LyAr00] *Lynch, John G.; Ariely, Dan*: Wine online: Search cost affect competition on price, quality, and distribution. In: *Marketing Science* 19 (2000) 1, S. 83–103.
- [MaWG95] *Mas-Colell, Andreu; Whinston, Michael D.; Green, Jerry R.*: Microeconomic Theory. Oxford University Press, USA 1995.
- [MeMa01] *Meredith, Lindsay; Maki, Dennis*: Product cannibalization and the role of prices. In: *Applied Economics* 33 (2001) 14, S. 1785–1793.
- [MeRS03] *Mehta, Nitin; Rajiv, Surendra; Srinivasan, Kannan*: Price Uncertainty and Consumer Search: A Structural Model of Consideration Set Formation. In: *Marketing Science* 22 (2003) 1, S. 58–84.
- [MoZS01] *Morton, Fiona Scott; Zettelmeyer, Florian; Silva-Risso, Jorge*: Internet Car Retailing. In: *Journal of Industrial Economics* 49 (2001) 4, S. 501–519.
- [Odly96] *Odlyzko, Andrew*: The Bumpy Road of Electronic Commerce. In: *Proceedings of WebNet 96, AACE* (1996), S. 378–389.
- [PaRS02] *Pan, Xing; Ratchford, Brian T.; Shankar, Venkatesh*: Can price dispersion in online markets be explained by differences in e-tailer service quality? In: *Journal of the Academy of Marketing Science* 30 (2002) 4, S. 433–445.
- [PaSR02] *Pan, Xing; Shankar, Venkatesh; Ratchford, Brian T.*: Price Competition Between Pure Play vs. Bricks-and-Clicks e-Tailers: Analytical Model and Empirical Analysis. In: *Baye, M. R.* (Hrsg.): *Advances in Applied Microeconomics: The Economics of the Internet and E-Commerce* 11 (2002), S. 26–61.
- [ReMc02] *Reardon, James; McCorkle, Denny E.*: A consumer model for channel switching behavior. In: *International Journal of Retail and Distribution Management* 30 (2002) 4, S. 179–185.
- [Salo79] *Salop, Steven C.*: Monopolistic Competition with Outside Goods. In: *The Bell Journal of Economics* 10 (1979) 1, S. 141–156.

- [ScGo02] *Schoenbachler, Denise D.; Gordon, Geoffrey L.*: Multi-channel shopping: Understanding what drives channel choice. In: *Journal of Consumer Marketing* 19 (2002) 1, S. 42–53.
- [ShRP99] *Shankar, Venkatesh; Rangaswamy, Arvind; Pusateri, Michael*: Customer Price Sensitivity and the Online Medium. *Smith Papers Online*. College Park, MD.
- [ShSR03] *Shankar, Venkatesh; Smith, Amy K.; Rangaswamy, Arvind*: Customer satisfaction and loyalty in online and offline environments. In: *International Journal of Research in Marketing* 20 (2003) 2, S. 153–175.
- [ShVa99] *Shapiro, C.; Varian, H.*: *Information Rules*. Harvard Business School Press 1999.
- [Simo92] *Simon, Hermann*: *Preismanagement: Analyse, Strategie, Umsetzung*. 2. Aufl., Gabler, Wiesbaden 1992.
- [Skie00] *Skiera, Bernd*: Preispolitik und Electronic Commerce - Preisdifferenzierung im Internet. In: *Wamser, C.* (Hrsg.): *Electronic Commerce*. München 2000, S. 117–130.
- [SmBr01] *Smith, Michael D.; Brynjolfsson, Erik*: Consumer Decision-making at an Internet Shopbot: Brand Still Matters. In: *Journal of Industrial Economics* 49 (2001) 4, S. 541–558.
- [SrRG05] *Srinivasan, Sundara Raghavan; Ramakrishnan, Sreeram; Grasman, Scott E.*: Identifying the effects of cannibalization on the product portfolio. In: *Marketing Intelligence and Planning* 23 (2005) 4/5, S. 359–371.
- [StBA02] *Steinfeld, Charles; Bouwman, Harry; Adelaar, Thomas*: The Dynamics of Click-and-Mortar Electronic Commerce: Opportunities and Management Strategies. In: *International Journal of Electronic Commerce* 7 (2002) 1, S. 99–119.
- [Stig61] *Stigler, George J.*: The Economics of Information. In: *The Journal of Political Economy* 69 (1961) 3, S. 213–225.
- [StTe02] *Stremersch, Stefan; Tellis, Gerard J.*: Strategic bundling of products and prices: A new synthesis for marketing. In: *Journal of Marketing* 66 (2002) 1, S. 55–72.

- [SuTh04] *Sullivan, Ursula Y.; Thomas, Jacquelyn S.*: Customer Migration: An Empirical Investigation Across Multiple Channels. http://www.business.uiuc.edu/Working_Papers/papers/04-0112.pdf.
- [Tern01] *Terna, P.*: Creating Artificial Worlds: A Note on Sugarscape and Two Comments. In: *Journal of Artificial Societies and Social Simulation* 4 (2001) 2.
- [TsAg04] *Tsay, Andy A.; Agrawal, Narendra*: Channel Conflict and Coordination in the E-Commerce Age. In: *Production and Operations Management* 13 (2004) 1, S. 93–110.
- [Vari95] *Varian, Hal R.*: Pricing Information Goods. University of Michigan (1995).
- [WaGJ04] *Wallace, David W.; Giese, Joan L.; Johnson, Jean L.*: Customer retailer loyalty in the context of multiple channel strategies. In: *Journal of Retailing* 80 (2004) 4, S. 249–263.
- [Ward01] *Ward, Michael R.*: Will Online Shopping Compete more with Traditional Retailing or Catalog Shopping? In: *Netnomics* 3 (2001) 2, S. 103–117.
- [Wern94] *Wernerfelt, Birger*: Selling Formats for Search Goods. In: *Marketing Science* 13 (1994) 3, S. 298–309.
- [Wik05] *Wikström, Solveig*: From E-Channel to Channel Mix and Channel Integration. In: *Journal of Marketing Management* 21 (2005), S. 725–753.
- [ZeMS06] *Zettelmeyer, Florian; Morton, Fiona Scott; Silva-Risso, Jorge*: How the Internet Lowers Prices: Evidence from Matched Survey and Automobile Transaction Data. In: *Journal of Marketing Research* 43 (2006) 2, S. 168–181.
- [Zett00] *Zettelmeyer, Florian*: Expanding to the Internet: Pricing and Communications Strategies When Firms Compete on Multiple Channels. In: *Journal of Marketing Research* 37 (2000) 3, S. 292–308.
- [ZwRL⁺03] *Zwick, Rami; Rapoport, Amnon; Lo, Alison King Chung et al.*: Consumer Sequential Search: Not Enough or Too Much? In: *Marketing Science* 22 (2003) 4, S. 503–519.

Ubiquitous Computing im Krankenhaus

Eine fallstudienbasierte Untersuchung betriebswirtschaftlicher Potenziale

Dipl.-Kfm. Falk Zwicker

Betreuer: Prof. Dr. Torsten Eymann
Lehrstuhl für Wirtschaftsinformatik (BWL VII)
Universität Bayreuth, 95440 Bayreuth
falk.zwicker@uni-bayreuth.de

Abstract

Die Vision des Ubiquitous Computing verspricht zukünftig die Allgegenwärtigkeit von Informationsverarbeitung. Das im Jahre 1991 von Mark Weiser formulierte Leitbild geht von einer immer höheren Anzahl von Computern bzw. Rechenleistung in der Umwelt aus, die fast unbemerkt im Hintergrund agieren und dem Menschen das Leben erleichtern. Im Rahmen des Ubiquitous Computing organisierte Informationssysteme sollen zukünftig helfen, schlecht strukturierte Probleme besser zu bearbeiten und den wirtschaftlichen Nutzen für Anwender zu erhöhen.

Gerade in Krankenhäusern treten im Rahmen ihrer individuellen Dienstleistungsproduktion diese Art von Problemen gehäuft auf. Zukünftig könnte Ubiquitous Computing, das derzeit vor allem mit Hilfe von Radio Frequency Identification-Technologie (RFID) umgesetzt wird, die Anforderungen für Flexibilitätsverbesserungen, Effizienzsteigerungen und verbesserte Prozesskoordination erhöhen. Techniken wie RFID ermöglichen die weitestgehend automatisierte Erfassung von Daten aus der Umwelt, die in einem kontinuierlichen Strom im Hintergrund arbeitenden Informationssystemen verfügbar gemacht werden.

Im Rahmen der angestrebten Arbeit soll ein theoretischer Bezugsrahmen zur Bewertung der ökonomischen Auswirkungen eines Einsatzes von Ubiquitous Computing zur Lösung bei schlecht strukturierten Problemen entwickelt und mithilfe von Fallstudien anhand eines praktischen Einsatzes im Gesundheitswesen getestet werden.

1 Einführung

1.1 Ausgangssituation und Motivation der Arbeit

Schlecht strukturierte Problemstellungen lassen sich in zahlreichen Institutionen und Organisationen vorfinden, deren Lösung eine große Herausforderung in der betriebswirtschaftlichen Forschung darstellt. Insbesondere die Dienstleistungsproduktion im Krankenhaus lässt sich unter diese Art der Problemstellung subsumieren. Erkrankungen von Patienten benötigen eine individuell abgestimmte Behandlung. Trotz großer Anstrengungen sind klinische Behandlungsabläufe daher nur schwer planbar, was wiederum erhebliche Einflüsse auf die Bereitstellung und den Einsatz von Ressourcen hat. Wollen Krankenhäuser ihre Wirtschaftlichkeit verbessern, müssen zukünftig auch diese Probleme zufrieden stellender gelöst werden.

Nicht nur die Besonderheiten im Rahmen der Leistungserstellung machen eine Optimierung beim Umgang mit schlecht strukturierten Problemen notwendig. Wandelnde Umweltbedingungen von Krankenhäusern erhöhen den Handlungsdruck. Die Umstellung auf ein Entgeltsystem mit fallbasierten Abrechnungspauschalen verstärkt die Bedeutung marktwirtschaftlicher Anreizstrukturen. Krankenhäuser sind angehalten die Effizienz zu steigern und dabei gleichzeitig die Versorgungsqualität zu sichern. Zudem gestaltet sich die Leistungserstellung wegen des medizinischen Fortschritts immer komplexer. Hinzu kommt eine stärkere Spezialisierung und Aufgliederung einzelner Fachbereiche. Krankenhäuser müssen insgesamt flexibler und reaktionsfähiger werden [HWAB02]. Neue Entwicklungen im Bereich der Informations- und Kommunikationstechnologie sollen dabei unterstützen und neue Wege aufzeigen, diese Probleme zu lösen.

Einen Beitrag dazu verspricht die technische Vision des „Ubiquitous Computing“. Darunter versteht man die Realisierung allgegenwärtiger Informationsverarbeitung, in der Rechner fast unsichtbar im Hintergrund arbeiten und als „smarte Objekte“ miteinander kommunizieren [Weis91]. Die mit dem Ubiquitous Computing organisierten Informationssysteme sollen die notwendigen Anforderungen von Krankenhäusern für Flexibilitätsverbesserungen, Effizienzsteigerungen und eine verbesserte Prozesskoordination erfüllen und so zur Lösung schlecht strukturierten Problemstellungen beitragen [MKSE03].

Radio-Frequency-Identification-Systeme (RFID-Systeme) stellen einen technischen Treiber des Ubiquitous Computing dar [FIDi03]. RFID ermöglicht die weitgehend automatisierte und zum

Teil autonome Erfassung von Umweltdaten und die Interaktion mit anderen informationsverarbeitenden Systemen.

Der Einsatz von RFID setzt voraus, dass die Kliniken Chancen und Risiken dieser Technologie korrekt einschätzen und bei ihren strategischen Überlegungen sowie Investitionsentscheidungen berücksichtigen. Zahlreiche Beispiele aus der Vergangenheit zeigen, dass falsche Prognosen bei der Implementierung von Informationstechnologien zu Fehlinvestitionen mit hohen Folgekosten geführt haben [SwRa04].

1.2 Zielsetzung der Arbeit

In der Arbeit sollen vor dem Hintergrund der beschriebenen Probleme die betriebswirtschaftlichen Potenziale eines im Rahmen des Ubiquitous Computing organisierten klinischen Informationssystems untersucht werden. Grundlage der Untersuchung ist ein Propositionengerüst, das mithilfe von Fallstudien in ausgewählten Kliniken getestet wird. Die Ergebnisse dienen als Basis für eine spätere Hypothesen- bzw. Theoriebildung und deren empirische Überprüfung.

Das Forschungsgebiet des Ubiquitous Computing ist breit gefächert, wobei vor allem technische Fragestellungen dominieren. Im Rahmen wirtschaftswissenschaftlicher Untersuchungen existiert derzeit wenig Erfolg versprechendes. Die meisten Arbeiten konzentrieren sich auf die Beschreibung des Machbaren, insbesondere im Rahmen einzelner Logistikprojekte. Diese sind wegen ihres zumeist speziellen Untersuchungsansatzes nur in begrenztem Umfang generalisierbar. Einen allgemein anerkannten theoretischen Bezugsrahmen zur Bewertung der ökonomischen Auswirkungen des Ubiquitous Computing auf schlecht strukturierte Problemstellungen in Organisationen gibt es zurzeit nicht.

Mit dieser Arbeit soll ein solcher Bezugsrahmen entwickelt werden, der durch einen praktischen Einsatz im Gesundheitswesen getestet wird. Wie noch zu zeigen, eignet sich die Dienstleistungsproduktion im klinischen Umfeld gut als Untersuchungsgegenstand. Aufgrund der individuellen Situation jedes einzelnen Patienten, existieren im Krankenhaus vor allem schlecht strukturierte Problemstellungen mit entsprechenden Auswirkungen auf Prozessabläufe. Mögliche Auswirkungen auf die Koordination und Integration der Leistungserstellungsprozesse in den Krankenhäusern werden dabei untersucht, um Wechselwirkungen von organisatorischem Wandel und Gestaltung der Informationssysteme zu erklären. Ferner werden die Motive der einzelnen Beteiligten für innovatives Verhalten erforscht.

2 Konzeptionelle Grundlagen

2.1 Die Vision des Ubiquitous Computing

Die Idee des Ubiquitous Computing entstand Ende der 1980er Jahre in den Xerox-Laboratorien des Palo Alto Research Center (PARC). 1988 wurde dazu ein eigenes Programm gestartet, in der das Ziel verfolgt wurde, die Defizite bei der Benutzung von Personal Computers (PCs) zu erforschen [WEGB99]. Doch schon bald wurden im Rahmen des Projekts neue Vorstellungen entwickelt, die in die Idee des Ubiquitous Computing einfließen. Das heutige Verständnis beruht im Wesentlichen auf der 1991 von Mark Weiser formulierten Vision. Dieser verband mit dem Begriff des Ubiquitous Computing die Vorstellung allgegenwärtiger Computer, die nahezu unsichtbar und unaufdringlich Menschen bei ihrem Handeln unterstützen [Weis91].

Weiser beschreibt in erster Linie keine neue Technologie der Informationsverarbeitung, sondern insbesondere eine neuartige Form der Organisation zur Anordnung und Vernetzung von Rechnern bzw. Rechenleistung in Informationssystemen [Eyma03]. Die Definition von Ubiquitous Computing erfolgt vor allem über die Beschreibung organisatorischer Charakteristika und nicht ausschließlich über technische Parameter.

Bei der Nutzung von Informationssystemen wird Ubiquitous Computing ein nachhaltiges Veränderungspotenzial der bisherigen Einsatz- und Verwendungsmöglichkeiten zugesprochen [Matt03]. Die Bedienung der Technik tritt dabei weitgehend in den Hintergrund. Funktion und Fähigkeit zur Informationsverarbeitung gehen dadurch jedoch nicht verloren, sondern werden für den Nutzer leichter verfügbar. In der Zukunft soll so ein „Internet der Dinge“ entstehen, in dem nicht wie bislang nur einzelne stationäre Computer, sondern Alltagsgegenstände mit Hilfe integrierter Computer untereinander vernetzt sind [Matt05].

Für das Ubiquitous Computing existiert in der Wissenschaft keine allgemein verwendete Definition. Der Begriff wird in zahlreichen Abwandlungen benutzt, die zumeist auf den Ausführungen Weisers basieren. Zudem haben sich neue Begriffe wie die Ambient Intelligence oder Pervasive Computing etabliert. Im Rahmen dieser Arbeit werden verschiedene Auslegungen untersucht und eine Definition des Ubiquitous Computing abgeleitet.

2.1.1 Ziele und Charakteristika des Ubiquitous Computing

Um eine Definition des Ubiquitous Computing zu erarbeiten, werden nachfolgend die Charakteristika dieser Organisationsform herausgestellt. Mark Weisers Vision basiert auf den bekannten Prinzipien und Techniken der Informationsverarbeitung, die sich seit einiger Zeit

durch technischen Fortschritt nach und nach verwirklichen lassen. Gemäß seiner Vision sollen beliebige Dinge des alltäglichen Lebens mit integrierten Prozessoren, Sensoren und Funkmodulen ausgerüstet werden. Das Ubiquitous Computing stellt mit Hilfe der Sensorik und anschließenden Verarbeitung ein detailliertes Abbild der physischen Welt in einem Informationssystem zur Verfügung. Mit den so gewonnenen Daten können mithilfe von Verarbeitung die Ziele des Ubiquitous Computing verwirklicht werden.

Eine Zielsetzung liegt in der Optimierung wirtschaftlicher Prozesse, die vor allem durch eine durchgängige Automatisierung erreicht werden soll [Matt05]. Dazu verhelfen soll die permanente Überwachung von Umgebungszuständen, um das hinterlegte Informationssystem bzw. dessen Nutzer mit aktuellen Daten zu versorgen, die dazu klassifiziert und strukturiert werden müssen, um die Versorgung mit den jeweils notwendigen Daten zu garantieren. Im Rahmen der angestrebten durchgängigen Automatisierung sollen Ubiquitous Computing-Systeme in der Lage sein, sich durch Kontexterkenkung situationsangepasst zu verhalten.

Die Ziele des Ubiquitous Computing sind nur durch eine umfassende Informatisierung der Umwelt zu erreichen. Diese ist durch eine hohe Integration vieler miniaturisierte Prozessoren, Sensoren und Funkmodule gekennzeichnet, welche in Dinge, Räume und Umgebungen eingebettet werden. Dahinter liegt ein Netzwerk als Infrastruktur, das die Kommunikation der einzelnen Rechner ermöglicht. Jede einzelne Recheneinheit im Ubiquitous Computing besitzt eine eindeutige Identität und ist in der Lage autonom mit anderen Dingen oder Menschen zu kommunizieren. Über die Sensorik ist es dem Informationssystem möglich, Daten über seine Umwelt zu erfassen und mittels Kontexterkenkung zu interpretieren. Ein weiteres wesentliches Merkmal ist die Lokalisierung der einzelnen Dinge [Flei01].

Derzeit lassen sich zwei Trends identifizieren: Zum einen entwickeln sich Sensornetze, die Phänomene in ihrer Umwelt erfassen (Daten heben) und in das Informationssystem einspeisen. Durch diese automatisierte Erfassung, die ohne Medienbrüche möglich wird, können wesentlich mehr und genauere Daten als bei manueller Eingabe durch Menschen erfasst werden. Das Informationssystem ist so in der Lage, ein umfangreicheres Bild der physischen Welt auf logischer Ebene darzustellen und für Rechenoperationen zu nutzen. Zum anderen werden die Dinge durch Ubiquitous Computing smart. Darunter ist zu verstehen, dass sie in der Lage sind, mit anderen Dingen oder dem Menschen zu kommunizieren. Grundlage dieser Kommunikation sind die in den Sensornetzen gewonnenen Daten und daraus abgeleitete Informationen. Smarte Dinge tragen meist eigene Software sowie Informationen in sich. Mit diesen Regeln wird ihr

Verhalten gesteuert. Durch die Vernetzung ist es eher von untergeordneter Bedeutung, wo Daten gespeichert sind. Entscheidend ist, dass diese im gesamten System verfügbar gemacht werden.

2.1.2 RFID als technischer Treiber des Ubiquitous Computing

Erste Ubiquitous Computing-Systeme sind in Pilotanwendungen in den letzten Jahren implementiert worden. Weiser war 1991 mit der Formulierung seiner Vision den damaligen technischen Möglichkeiten um Jahre voraus. Technisch ist es seit einiger Zeit jedoch möglich, erste Anwendungen und Informationssysteme, die in der Form des Ubiquitous Computing organisiert sind, zu etablieren. Als ein Treiber hat sich dabei die schon seit längerem bekannte Technologie der Radio Frequency Identification (RFID) erwiesen, die momentan in vielen Bereichen Einzug hält [Matt05]. RFID-Systeme können dem Themenkomplex des Ubiquitous Computing zugeordnet werden, wenn die gewonnenen Daten in ein dahinter liegendes Informationssystem gespeist werden.

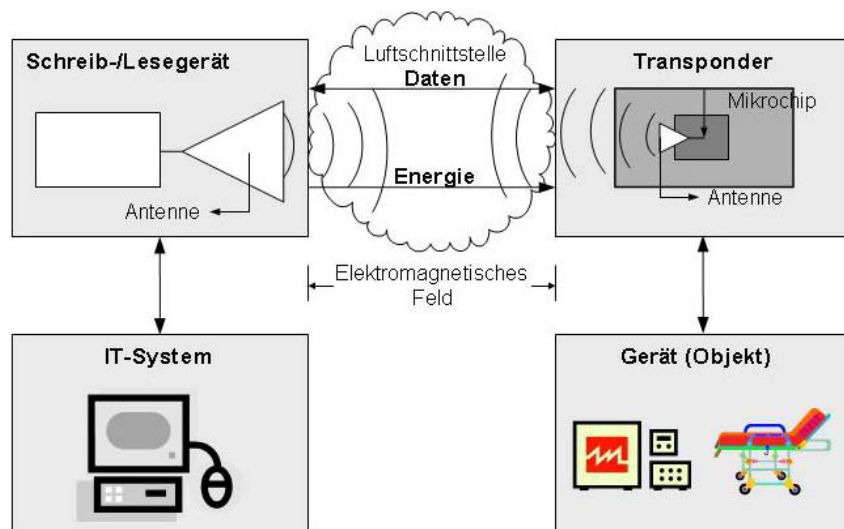


Abb. 1: Funktionsweise von RFID-Systemen [ZwRE06].

Die RFID-Technologie wird dazu genutzt, Objekte kontaktlos über eine Luftschnittstelle zu identifizieren. Technisch gesehen werden dazu Transponder verwendet, die mit einem Hochfrequenzsignal bestrahlt werden. Dieses Signal wird decodiert und ein Antwortsignal zurückgesendet. Häufig besteht es aus einer eindeutigen Identifikationsnummer, die mit Informationen aus einer Datenbank verknüpft wird. Die Energiegewinnung der Transponder

erfolgt entweder aus einer eigenen Quelle oder speist sich aus dem Signal, mit dem sie bestrahlt werden. Aktuelle Transponder sind in der Lage, innerhalb von Sekundenbruchteilen einige hundert Bits Daten zu übertragen. Wenn sie über einen eigenen Speicher verfügen, können sie sogar beschrieben werden. Auch die Koppelung an Sensoren, beispielsweise zur Temperaturermittlung, ist möglich. Transponder bestehen aus einem meist wenige Quadratmillimeter großem Chip und einer Antenne aus sehr dünnem Kupferkabel. Die Chips sind in der Regel sehr robust und eignen sich für viele Einsatzbereiche. Probleme beim Schreib-/Leseprozess bestehen derzeit allerdings noch bei Flüssigkeiten und in metallischen Umgebungen.¹

2.2 Die Dienstleistungsproduktion im Krankenhaus

2.2.1 Leistungserstellung im Krankenhaus als Produktion von Dienstleistung

Als Kernleistung eines Krankenhauses lässt sich der Behandlungs- bzw. Versorgungsprozess von Patienten definieren. Dieser dient dem Ziel, den Gesundheitszustand von Patienten wiederherzustellen oder zumindest zu verbessern [Lang97]. Er reicht von der Aufnahme des Patienten über die Diagnostik, therapeutische Maßnahmen bis zu seiner Entlassung. Leistungen des Krankenhauses werden individuell an Patienten vollbracht und lassen sich den Dienstleistungen zuzuordnen. Dienstleistungen könne durch konstitutive Merkmale charakterisiert und eingrenzt werden. Ihre Erstellung orientiert sich an drei Dimensionen, der Potenzial-, der Prozess- und der Ergebnisdimension [Sibb04]:

- Um eine Leistung erzeugen zu können, müssen zunächst aus internen Einsatzfaktoren Leistungspotenziale bereitgestellt werden.
- Die Erstellung der Dienstleistung benötigt einen externen Faktor, der durch den Kunden – im Fall von Krankenhäusern durch den Patienten – zur Verfügung gestellt wird.
- Leistungsergebnisse von Dienstleistungen sind immateriell.

¹ Für weitergehende Informationen zur RFID-Technologie sei an dieser Stelle auf die Ausführungen von Finkenzeller verwiesen [Fink02].

Die Nachfolgende Graphik veranschaulicht den mehrstufigen Dienstleistungsprozess eines Krankenhauses, wobei der zu behandelnde Patient als externer Faktor definiert ist:

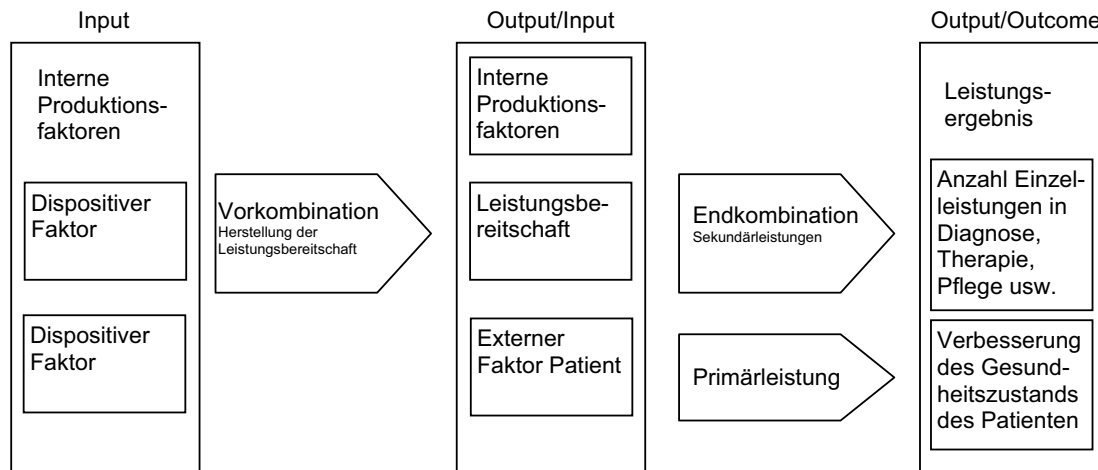


Abb. 2: Mehrstufige Dienstleistungsproduktion im Krankenhaus [Sibb04].

Um überhaupt Dienstleistungen zu produzieren, müssen Krankenhäuser im Rahmen einer Vorkombination (Input) Leistungspotenziale bereitstellen. Dazu zählen die Fähigkeiten zur Behandlung und der weiteren Versorgung von Patienten. Diese können erst durch das Hinzutreten des externen Faktors Patient genutzt werden. Ohne die Integration dieses externen Faktors ist eine Produktion der Dienstleistung nicht möglich. Wie bei den meisten Dienstleistungen, fallen die Leistungserstellung und deren Abgabe im Krankenhaus zeitlich zusammen. Dies zieht hohe Anforderungen an die Flexibilität der Produktionsfaktoren nach sich und ermöglicht keinen oder einen nur sehr kurzen Planungshorizont.

2.2.2 Dienstleistungen im Krankenhaus – individuelle und schlecht strukturierte Probleme

Durch ihre Individualität sind Krankenhausprozesse dynamisch und offen. Jederzeit können Änderungen im vorgesehenen Ablauf auftreten. Alle Dienstleistungen im Krankenhaus sind in hohem Maße integrativ und indeterminiert. Prägend für diese Prozesse sind somit Probleme bei der Kapazitätsanpassung- und Auslastung. Ihr hohes Maß an Individualität beschränkt die Möglichkeiten zur Standardisierung erheblich [Seel93]. Zwar weisen einzelne Teilprozesse von Krankenhausdienstleistungen wenig Unterschiede auf und sind damit durchaus bis zu einem gewissen Maße standardisierbar, jedoch bleibt für die Gesamtleistung immer noch ein

hoher individueller Anteil, der an den jeweiligen Patienten (externer Faktor) angepasst werden muss.²

Dienstleistungsprozesse im Krankenhaus sind neben ihrer Individualität sehr wissensintensiv. Je mehr Informationen verfügbar gemacht werden, desto besser kann auf die jeweiligen Bedürfnisse von Patienten reagiert und eingegangen werden. Informations- und Kommunikationstechnologien kommt dabei eine wesentliche unterstützende Rolle bei der Dienstleistungsproduktion zu. Eine gezielte Steuerung und Nutzung des Informationsflusses ermöglicht die systematische Aufbereitung relevanter Daten.

Insgesamt können Behandlungsprozesse im Krankenhaus als schlecht strukturierte Aufgaben angesehen werden. Variierende Behandlungsdauern, abrupte Änderungen der Behandlung – etwa bei Komplikationen – erwartete Krankheitsverläufe und äußere Störungen, die z.B. durch Notfälle auftreten, sind nur einige Faktoren, die eine vollständige Abbildung des Behandlungsprozesse verhindern und eine ständige Neuanpassung erfordern.

2.3 Relevante Forschungsbereiche

Verschiedene Forschungseinrichtungen beschäftigen sich mit ökonomischen Fragestellungen im Rahmen von Ubiquitous Computing. Die meisten Untersuchungen basieren zumeist auf Laborergebnissen. Eine Evaluation unter realen Bedingungen hat aufgrund der mangelnden Verbreitung häufig noch nicht stattgefunden. Einzig im Logistikbereich, bei dem bereits zahlreiche RFID-Lösungen implementiert wurden, scheint man inzwischen über genügend Daten für eine ökonomische Bewertung zu verfügen. Andere Forschungsansätze, die auf ökonomische Bestimmungsfaktoren abzielen, fokussieren sich auf die Identifizierung von Nutzenpotenzialen und ökonomische Bestimmungsfaktoren.

Ferner existieren einige wenige Konferenzpapiere in Form deskriptiver Fallstudien, die sich mit dem Healthcare-Bereich und damit dem Problemfeld schlecht strukturierter Probleme befassen. Zumeist handelt es sich um reine Darstellungen von Pilotprojekten. Die wirtschaftlichen Potenziale werden nur unzureichend untersucht und teilweise aus Managementmethoden abgeleitet.

² So sind bestimmte Eingriffe einander sehr ähnlich. Z.B. verläuft eine Operation am Blinddarm hinsichtlich Technik und Ablauf meist standardisiert. Allerdings variieren die Operationsvor- und Nachbereitung bei unterschiedlichen Patienten erheblich. Zu nennen wären beispielsweise besondere Maßnahmen aufgrund individueller Vorerkrankungen oder Allergien, wodurch der Gesamtprozess einen hochgradig individuellen Charakter besitzt.

3 Inhaltliche und methodische Vorgehensweise

Für die geplante Untersuchung wird die Methode der explorativen Fallstudie angewandt [Ried06], mit der Erfahrungen ausgewählter Krankenhäuser analysiert werden, die RFID-Lösungen implementiert haben. Die Datenerhebung erfolgt durch semistrukturierte Experteninterviews [Hopf95]. Nachfolgend wird die verwendete Methode wissenschaftlich eingeordnet und das inhaltliche sowie methodische Vorgehen im Rahmen der Dissertation beschrieben.

3.1 Inhaltliche Vorgehensweise

Die Arbeit gliedert sich in fünf Teile. Nach einer Einführung werden im zweiten Teil konzeptionelle Grundlagen erarbeitet. Diese beinhalten zunächst eine Diskussion über das Ubiquitous Computing. Von Weisers Vision ausgehend, werden die Charakteristika und Ziele dieser Organisationsform von Informationstechnologie näher erörtert. Die RFID-Technologie als ein technischer Treiber des Ubiquitous Computing wird vorgestellt und in die zu behandelnde Thematik eingeordnet.

Anschließend werden die Merkmale der Leistungserstellung im Krankenhaus als Form betriebswirtschaftlich erfassbarer Dienstleistung dargestellt und somit der Untersuchungsraum näher betrachtet. Dabei werden Prozessmerkmale herausgearbeitet und mithilfe Organisationstheorie als schlecht strukturierte Probleme eingeordnet.

Anschließend werden die für die Untersuchung relevanten Forschungsbereiche dargestellt. Um die Wechselwirkung von organisatorischem Wandel und Veränderungen von Informationssystemen zu erklären, werden die Erkenntnisse der Koordinationstheorie erörtert. Diese dienen in den weiteren Überlegungen als theoretisches Fundament [MaCr91]. Des Weiteren werden relevante Aspekte der Adaption von IT-Innovationen untersucht [BuRo81]. Die theoretischen Konstrukte entlehnen sich zumeist aus der Organisationstheorie und ausgesuchten Modellen der Wirtschaftsinformatik.

Im dritten Teil werden auf dieser Basis Propositionen entwickelt, die die Auswirkungen von Ubiquitous Computing auf Prozessabläufe in den Krankenhäusern erklären. Sie dienen gleichzeitig als Grundlage für die Experteninterviews für die Fallstudien [Ried06]. Eine genaue Einordnung der Methodik und der Propositionen folgen im nächsten Abschnitt.

Vor Durchführung der Fallstudien werden bei einer Vorstudie mehrere Fachleute befragt, um die Relevanz der aufgestellten Hypothesen empirisch zu testen. Als Interviewpartner werden

Experten aus Institutionen gewählt, die sich mit der Implementierung von RFID-Lösungen befasst haben.

Im vierten Teil erfolgt die Durchführung der Fallstudien. Zur Datenerhebung werden Projektverantwortliche und IT-Leiter aus Krankenhäusern befragt. Die Ergebnisse werden im fünften Kapitel analysiert und interpretiert. Die in den Fallstudien empirisch getesteten Hypothesen liefern einen theoretischen Bezugsrahmen. Die Arbeit schließt mit einer Zusammenfassung.

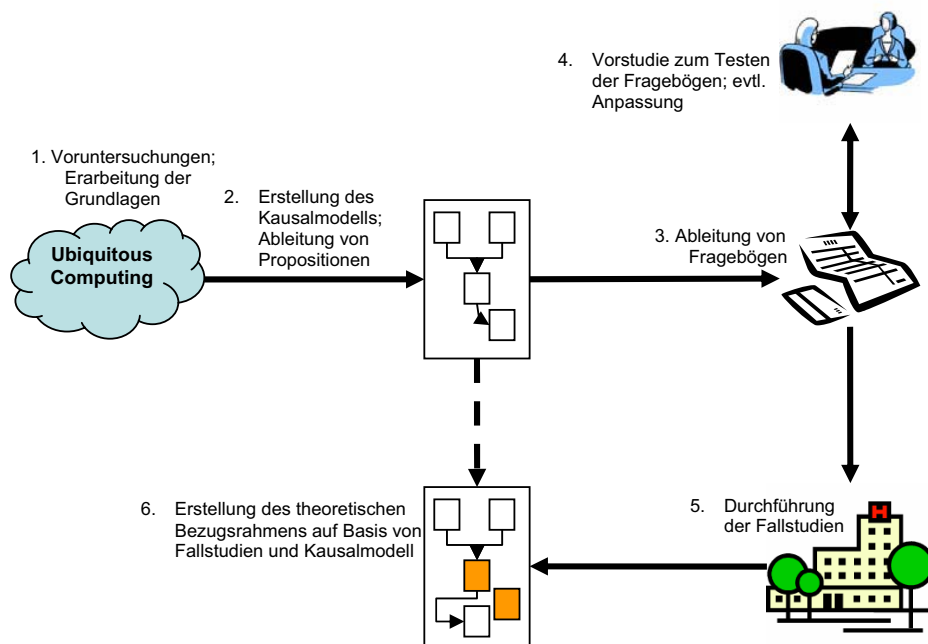


Abb. 3: Arbeitsschritte der geplanten Untersuchung.

3.2 Auswahl und Einordnung der Forschungsmethode

Mit dem Dissertationsvorhaben soll ein reales Phänomen im Schnittstellenbereich von Wirtschaftsinformatik und Betriebswirtschaftslehre untersucht werden. Als Ziel der Untersuchung steht ein Erkenntnisgewinn über die Frage, wie und warum der Einsatz von Ubiquitous Computing-Technologie in einer Umgebung mit zahlreichen unstrukturierten Problemen ökonomisch positiv wirken kann. Zur Durchführung der angestrebten Untersuchung wird eine erklärende Methode benötigt. Diese wird aus der Empirie gewählt, die ursprünglich aus den Sozialwissenschaften stammt. Empirismus kennzeichnet eine Strömung in der Wissenschaft, die davon ausgeht, dass alles Wissen über die Wirklichkeit aus Sinneserfahrung, also Wahrnehmung resultiert.

Die Arbeit strebt mit ihrer Untersuchung einen Beitrag zur Theoriebildung an. Theorien stellen zusammenhängende Reihen von singulären und universellen Aussagen dar, die es über ein reines Beschreiben hinaus ermöglichen, den Gegenstand einer Theorie zu begründen, zu erklären oder zu verstehen. Diese Verallgemeinerbarkeit wird durch Annahmen über Werte, Raum und Zeit begrenzt [Atte95].

Die Grundlage der Theoriebildung bilden Konstrukte und Propositionen, die einen theoretischen Bezugsrahmen formen. Konstrukte sind dabei als begriffliche Annäherungen an ein bestimmtes Phänomen zu verstehen. Propositionen spezifizieren die Beziehungen zwischen Konstrukten. Aus Propositionen lassen sich in einem späteren Schritt Hypothesen ableiten, die Beziehungen zwischen Variablen als messbare Größen spezifizieren. Variablen und Hypothesen gemeinsam konstituieren ein Modell bzw. eine Theorie.

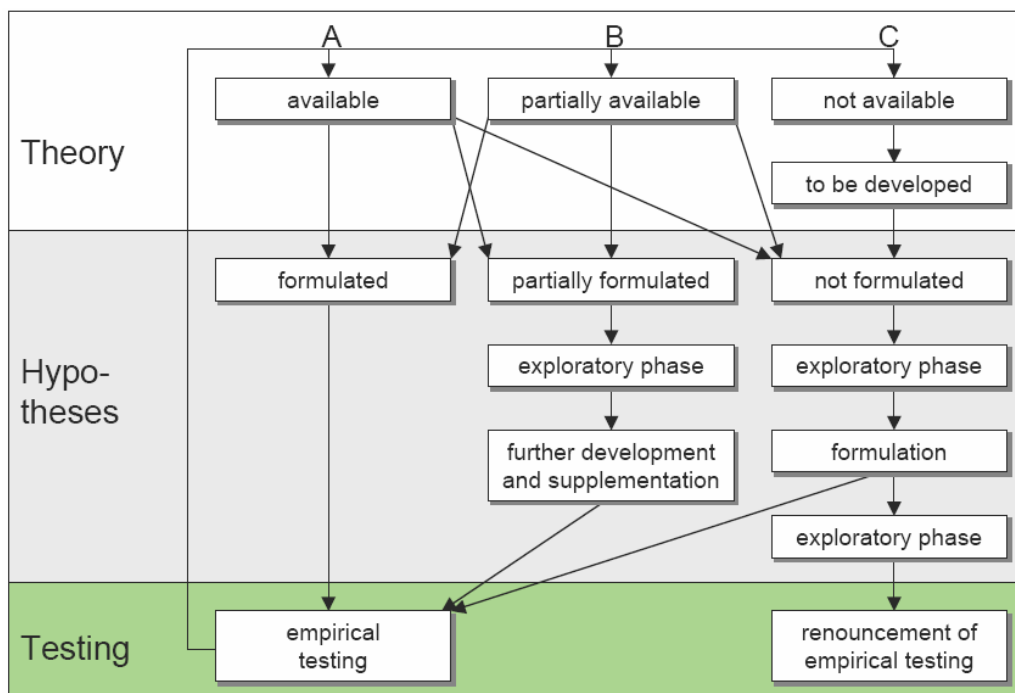


Abb. 4: Explorativer vs. konfirmatorischer Empirismus [Atte95].

Mit der Wahl einer empirischen Methode zielt die Arbeit auf ein systematisches Sammeln von Sinneswahrnehmungen bzw. Daten ab, um diese zu analysieren. Während im Bereich des konfirmatorischen Empirismus bekannte Zusammenhänge auf ihre Gültigkeit überprüft werden, sollen mit explorativen Untersuchungen neue, unbekannte Strukturen aufgedeckt werden. Aus den Ergebnissen wird dann ein Bezugsrahmen entwickelt, der als Basis für die spätere Hypothesenbildung und deren Test dient.

Aufgrund einer fehlenden wirtschaftswissenschaftlichen Theorie bzw. Hypothesen wird für die Arbeit ein exploratives Design gewählt. Beim gewählten Untersuchungsraum handelt es sich um eine komplexe Ausgangssituation, in der relativ wenig über den Untersuchungsgegenstand und die relevanten Variablen bekannt ist. Ferner stehen keine allgemein anerkannten Hypothesen zur Überprüfung zur Verfügung. Es sind zwar bestimmte Zusammenhänge zwischen Größen erkennbar, aber die Kausalität zwischen ihnen ist nicht genau zu erklären. Die Datenerhebung bei der gewählten explorativen Fallstudie erfolgt mit Hilfe semistrukturierter Experteninterviews.

3.3 Durchführung der Fallstudie und Entwicklung der Propositionen

Arbeiten, die sich explorativer Methodenelemente bedienen, stehen häufig in der Kritik, intersubjektiv nicht nachvollziehbar zu sein. Um dem Anspruch an wissenschaftliches Arbeiten zu genügen und den Verfasser selbst vor eigenen Fehlinterpretationen der Untersuchungsergebnisse zu schützen, muss sich die Arbeit nach einem strikten Ablauf richten und strengen Implikationen folgen. Die Durchführung der Fallstudien orientiert sich an folgendem Schema [Wron05]:

- Offenlegung theoretischen Vorwissens.
- Ableitung von Propositionen aus theoretischem Vorwissen.
- Auswahl von Untersuchungsfällen (Sampling).
- Entwicklung und Pretest des Interviewleitfadens.
- Durchführung der Fallstudie (Datenerhebung).
- Testen der Propositionen mit Hilfe der Fallstudien (Datenauswertung).
 - Fallbezogene Analyse.
 - Fallübergreifende Analyse.
- Überprüfung der Annahmen aus theoretischem Vorwissen.
- Ableitung eines theoretischen Bezugsrahmens als Ergebnis der Fallstudien und Vorbereitung der Hypothesenbildung.

Die gemachten Schlussfolgerungen müssen ausschließlich im Untersuchungsmaterial begründet und durch Dritte nachvollziehbar sein. Auch sollten die durchgeführten Interviews einem vorstrukturierten Ablauf folgen und die transkribierten Daten dem Interviewpartner zur inhaltlichen Zustimmung vorgelegt werden. Die im Vorfeld erstellten Propositionen werden mithilfe der aus den Interviews gewonnenen Daten so entweder bestätigt oder widerlegt. Als

Ergebnis der Arbeit wird ein überarbeiteter Propositionenrahmen formuliert, der aus Ausgangspunkt für weitere Untersuchungen dient.

4 Abgabe

Die Erstellung der Propositionen und die Erhebung der Daten sind für die ersten drei Quartale in 2007 geplant. Die Arbeit soll im ersten Quartal 2008 fertig gestellt sein.

Literaturverzeichnis

- [Atte95] *Atteslander, Peter*: Methoden der empirischen Sozialforschung. 8. Aufl. Gruyter, Berlin 1995.
- [BuRo81] *Bullen, Christine V.; Rockart, John F.*: A Primer on Critical Success Factors. MIT Center for Information Systems Research, Working Paper No. 69, Cambridge 1981.
- [Eyma03] *Eymann, Torsten*: Digitale Geschäftsagenten: Softwareagenten im Einsatz. Springer, Berlin 2003.
- [Fink02] *Finkenzeller, Klaus*: RFID-Handbuch. 3. Aufl. Hanser, München 2002.
- [Flei01] *Fleisch, Elgar*: Betriebswirtschaftliche Perspektiven des Ubiquitous Computing. In: *Bühl, H.; Huther, A.; Reitwiesner, B.*: Information Age Economy. Physica-Verlage, Heidelberg 2001, S. 177-191.
- [FIDi03] *Fleisch, Elgar; Dierkes, Markus*: Ubiquitous Computing aus betriebswirtschaftlicher Sicht. In: WIRTSCHAFTSINFORMATIK 45 (2003) 6, S. 611-620.
- [Hopf95] *Hopf, Christel*: Qualitative Interviews in der Sozialforschung – Ein Überblick. In: *Flick, Uwe (Hrsg.)*: Qualitative Sozialforschung. Rowohlt, Hamburg 2002, S. 177-182.

- [HWAB02] *Haux, Reinhold; Winter, Alfred; Ammenwerth, Elske; Birgl, Birgit*: Strategic Information Management in Hospitals. An Introduction to Hospital Informations Systems. Springer, New York 2002.
- [Lang97] *Lang, Herbert*: Erfolgsfaktoren privater Krankenanstalten. Josef Eul-Verlag, Köln 1997.
- [MaCr91] *Malone, Thomas W.; Crowston, Kevin*: Toward an Interdisciplinary Theory of Coordination. MIT Center for Coordination Sciences, Technical Report No. 120, Cambridge 1991.
- [Matt03] *Mattern, Friedemann*: Vom Verschwinden des Computers – Die Vision des Ubiquitous Computing, in: *Mattern, Friedemann*: Total vernetzt – Szenarien einer informatisierten Welt. Springer, Berlin 2003.
- [Matt05] *Mattern, Friedemann*: Die technische Basis für das Internet der dinge. In: *Fleisch, Elgar; Mattern, Friedemann*: Das Internet der dinge – Ubiquitous Computing und RFID in der Praxis. Springer, Berlin 2005, S. 39-66.
- [MKSE03] *Müller, Günter; Kreutzer, Michael; Strasser, Moritz; Eymann, Torsten; Hohl, Adolf; Nopper, Norbert; Sackmann, Stefan*: Geduldige Technologie für ungeduldige Patienten: Führt Ubiquitous Computing zu mehr Selbstbestimmung? In: *Mattern, Friedemann* (Hrsg.): Total vernetzt. Springer, Berlin 2003, S. 159-186.
- [Ried06] *Riedl, René.*: Erkenntnisfortschritt durch Forschungsfallstudien: Überlegungen am Beispiel der Wirtschaftsinformatik. In: *Zelewski, Stephan; Akca, Naciye* (Hrsg.): Fortschritt in den Wirtschaftswissenschaften: Wissenschaftstheoretische Grundlagen und exemplarische Anwendungen. Deutscher Universitäts-Verlag, Wiesbaden 2006, S. 113-145.
- [Seel93] *Seelos, H.-J.*: Die konstitutiven Merkmale der Krankenhausleistungsproduktion. In: *führen und wirtschaften im Krankenhaus*, 10. Jg. (1993), Nr. 2, S. 108-116.
- [Sibb04] *Sibbel, Rainer*: Produktion integrativer Dienstleistungen. Deutscher Universitätsverlag, Wiesbaden 2004.

- [SwRa04] *Swanson, E. Burton; Ramiller, Neil*: Innovating Mindfully with Information Technology. In: MIS Quarterly, Vol. 28 No. 4 (2004), S. 553-583.

- [WeGB99] *Weiser, Mark; Gold, R.; Brown, J.*: The origins of ubiquitous computing research at PARC in the late 1980s. In: IBM Systems Journal, Vol. 38 No. 4 (1991), S. 693-696.

- [Weis91] *Weiser, Mark*: The Computer for the 21st Century. In: Scientific American, Vol. 265 No. 3 (1991), S. 94-104.

- [Wron05] *Wrona, Thomas*: Die Fallstudienanalyse als wissenschaftliche Forschungsmethode. ESCP-EAP Working Paper Nr. 10, März 2005.

- [ZwRE06] *Zwicker, Falk; Reiher, Michael; Eymann, Torsten*: RFID im Krankenhaus - neue Wege in der Logistik, in: *Eymann, Torsten; Hampe, Felix; Koop, Andreas; Niemann, Christoph (Hrsg.)*: Mobiles Computing in der Medizin (Proceedings zum 6. Workshop der GMDS-Arbeitsgruppe "Mobiles Computing in der Medizin" am 2. Juni 2006 in Frankfurt/Main). Aachen 2006, S. 68-74.

Autoren

Anoraganingrum, Dwi, 247

Brandt, Christian, 233

Burghardt, Matthias, 99

Faupel, Titus, 219

Fiege, René, 11

Gille, Daniel, 29

Gruber, Gottfried, 317

Krabichler, Thomas, 187

Luckner, Stefan, 207

Müller, Claudia, 299

Niemann, Christoph, 49

Nowey, Thomas, 281

Piechocki, Maciej, 133

Reuter, Claudia, 63

Schlereth, Christian, 169

Slamka, Christian, 263

Streitberger, Werner, 83

Willenweber, Kim, 115

Wilde, Thomas, 151

Zwicker, Falk, 335

Bereits seit Anfang der 1990er Jahre wird jungen Wissenschaftlern im Vorfeld der Tagung "Wirtschaftsinformatik" ein Doctoral Consortium als unterstützendes Forum angeboten. Diese Einrichtung wurde auch zur größten internationalen Konferenz der Wirtschaftsinformatik, der WI 2007 in Karlsruhe fortgeführt. Dieser Band fasst die zum Vortrag ausgewählten Beiträge zusammen.

Since the early 1990es, young researchers participate in the doctoral consortium series, co-located with the "Wirtschaftsinformatik" conference. This volume contains the selected papers of 20 PhD candidates of the 2007 doctoral consortium in Karlsruhe.